数学方法论选讲

徐利治 著

朱梧槚、袁相碗 校

華中工学院出版社

内 容 简 介

本书用十来个典型的专题，对数学的发展规律和思想方法，进行了认真的研究和讨论，书中着重介绍了数学模型方法、公理化方法、映射反演原则、结构主义和伽罗瓦群的思想；分析了悖论与数学基础问题的关系以及对数学发展的影响；探讨了逻辑主义、直觉主义、形式主义等数学诸流派的观点、方法以及它们的成因；叙述了数学家在数学研究中的发现、发明与创新过程的心智状态，本书用辩证的观点，总结了历史上著名数学家希尔伯特等人的成长条件和成功的经验。

本书可作为理工科大学高年级学生和研究生的选修课教材，也可供数学工作者，哲学工作者以及教师们参考。

数学方法论选讲

徐利治　著

责任编辑　李立鹏

华中工学院出版社出版

（武昌喻家山）

湖北省新华书店发行　各地新华书店经售

咸宁县印刷厂印刷

开本：850×1168 1/32　印张：5.875　字数：150,000

1983年 4 月第一版　1983年 4 月第一次印刷

印数：1—15,000　统一书号：13255—013　定价：0.90元

前　言

“数学方法论”是一门很重要的学问，在概念上它理应属于科学方法论的一个特殊范畴。但本书并不是关于数学方法论的系统论述，而只是选择了十来个公认为比较有趣的专题，对它们作了介绍、分析和讨论。这些介绍或讨论，相信对从事数学或数理哲学的科研工作者和教师们会有一定的参考价值。

“尽信书不如无书”。建议读者要尽可能采取科学的分析态度来阅读这本书。对书中提到的许多著名数学家的观点和主张，自然不宜盲目崇拜，通盘接受。这好比消化任何食物那样，必须吸收其养料，排泄其糟粕。书中部分内容还包含着我和合作者们的一些论点或心得，这就更希望读者给予分析评论并不吝指正。

本书主要题材是根据作者近两年来在大连、长春、武汉三地讲课用的提纲和讲义充实形成的。南京大学、辽宁师院等高校数学系均将采用本书开设选修课程。如所知，这样一门课程在国内外尚未正式形成讲授科目，所以并无现成教材可资借鉴。因此，这里无论是题材内容的选择或是讲授方法，都不能认为是充分定型的。然而，使作者感到欣慰的是：不少数学专业和自然辩证法专业的高年级学生和研究生听了这门课之后，认为无论在数学的或哲学的思考中都能受到启发和得到提高。

应该提到的是，本书第3讲、第8讲与第9讲是根据作者和朱梧槚、袁相碗、王兴华、郑毓信等同志合作的论文改写而成的。

我要对我的学术助手田铁虹、高俊斌二同志和组合数学方向的研究生们表示谢意，因为他们曾耐心地帮助誊清了原稿。还应提到孙革同志，她曾分担了整理和抄写原稿的大量工作。最后，我特别要对华中工学院出版社将拙著纳入出版计划以及编辑部李立鹏等同志所付出的细致劳动表示诚挚感谢。

徐利治

1982年12月于大连

目 录

第1讲 数学方法论引论

§1 研究数学方法论的意义和目的

什么叫**方法论**？方法论(methodology)就是把某种共同的发展规律和研究方法作为讨论对象的一门学问。英文methodology一词又译为方法学。如所知，各门科学都有方法论，数学当然也有它自己的方法论。

数学方法论主要是研究和讨论数学的发展规律、数学的思想方法以及数学中的发现、发明与创新等法则的一门学问。

数学是一门工具性很强的科学，它和别的科学比较起来还具有较高的抽象性等特征，为了有效地发展它、改进它、应用它或者把它很好地传授给学生们，就需要对这门科学的发展规律、研究方法、发现与发明等法则有所掌握。因此，数学研究工作者、数学教师、科技工作者，以及高年级大学生、研究生等都需要知道一些数学方法论。

由于数学领域里的许多概念与理论题材都是通过人脑的抽象思维形式表现出来的，这里不仅包含有思维对象（数学本体）的辩证法，而且还有着思维运动过程（认识与反映过程）的辩证法。所以数学方法论还给哲学家、自然辩证法研究工作者以及心理学家们提供了值得分析研究的素材。凡是看过恩格斯《自然辩证法》的读者都知道，即使在初等数学里也充满着辩证法。

我们又知道，数学方法论中的许多方法和原理是从数学发展史中总结归纳出来的，所以数学工作者还必须学习一点数学史。

从近代数学发展史中，我们看到有许多杰出的数学家曾围绕

着数学基础问题展开了一系列争论，以致形成了各个著名的流派，如逻辑主义派、直觉主义派、形式主义派与柏拉图主义派等。直到现今，这些流派的观点主张对数学体系的内在发展，还继续产生着不同程度的影响。

各个数学流派对待数学基础问题的研究，各有其方法论主张。事实上，他们各有所偏，各有所见。只有运用科学的反映论观点，才能从他们的观点主张中分析总结出较为正确的数学方法论观点。因此，对于今日的数学工作者说来，无论是为了掌握、运用或者去发展数学方法论，都必须自觉地采取科学的反映论观点（即辩证法的反映论观点）去考察问题和分析问题。

§2　宏观的方法论与微观的方法论

数学科学的发展规律可以从数学发展史的丰富材料中归纳分析出来。由于数学史是人类社会科学技术发展史的一个组成部分，数学发展的巨大动力源泉与社会生产实践及技术发展的客观要求紧密相连，因此，数学发展规律的研究，如果撇开数学内在因素不提，那是属于**宏观的数学方法论**范畴。

数学工作者研究数学课题时，也可以不考虑数学发展的外在推动力，专就数学内部体系结构中的特定问题来进行分析研究。这样，就需要考虑采取最有效的数学研究方法，需要懂得数学发现与数学创造等各种法则。这些属于研究工作者个人必须遵循的方法与法则的研究，可以称之为**微观的数学方法论**。

看来，历史上最卓越的数学家如牛顿、欧拉、高斯、傅立叶、拉普拉斯等人，既精通微观的数学方法论，也懂得宏观的数学方法论。否则，他们的成就与贡献不可能对社会生产技术的发展产生那样深远的影响。一般说来，凡是具有历史眼光的数学家，他们的贡献成果，往往起着承上启下的作用，因而总是带有经久不灭的光辉。怎样才能获得“历史眼光”呢？这就需要通过

数学史的研究去理解一些宏观数学方法论的基本知识。

这里值得介绍的是，美国数学教授M.Kline曾在1972年出版了一本厚达1200页的巨著——《古今数学思想》，系统地叙述和总结了古今数学思想发展史。该书包藏大量的题材，可作为我们研究数学方法论的一本宝贵的参考资料。还有E.T.Bell的一本名著《数学人物》（1937年出版，1965年重版），其中翔实地记录了古今30多位杰出数学家的生活经历与工作历史，也很有参考价值。

数学家成长规律的一般分析，显然也应属于宏观的方法论；但本书只着重讨论微观的数学方法论，所以仅借用希尔伯特成功的典型范例来描绘一下关于数学人才成长的社会因素的作用。

§3　略论希尔伯特成功的社会因素

分析一位杰出数学家成功的社会因素，对于正在成长着的青年数学工作者和从事数学教育的数学教师们说来，都会得到有益的启发。这种分析至少对消除"天才自成"的糊涂思想，会起到一定的作用。

我们选择希尔伯特（David Hilbert，1862-1943）这个例子，因为这位数学家的成长、发展和获得巨大成功的经历已经成为现代人才学上的一个典型例子。

按照历史唯物主义的观点来看，"天才人物"都是社会的产物。他们只有适应时代的要求，回答和解决历史进程中出现的重大问题，才能取得成功。分析任何一个天才人物成功的因素时，应当象恩格斯那样，不能离开当时社会历史条件和文化发展条件以及反映这些条件的时代要求。所以对希尔伯特的分析，也应遵循这样的原则。本节内容主要取材于《大连工学院自然辩证法通讯》上刘永振同志的一篇文章，该文显然是参考了希尔伯特传记写成的。

希尔伯特出身于东普鲁士的古都哥尼斯堡（现名加里宁格勒）。中、青年时代，他曾对代数不变式论、代数数论、几何基础等科目作出了重要贡献。中年以后，他发展了变分法、积分方程、函数空间理论、数学物理方法、数理逻辑及证明论等数学分支。1899年出版的一本名著《几何学基础》，成为近代公理化方法的代表作，且由此推动形成了“数学公理化学派”。所以，希尔伯特是近代形式公理学派的创始人。1900年，年届38岁，他在国际数学会议上以卓越的远见和洞察力提出了数学上未解决的23个难题，即有名的“希尔伯特问题”，推动了半个多世纪以来各个数学分支的发展。

我们知道，十九世纪七十年代初，德国实现了统一，经济空前高涨，成为世界科学活动的主要中心，这是产生一大批德国数学家的主要社会原因。高水平的一群人才中，必有出类拔萃者。希尔伯特生逢盛世，成为出类拔萃者，这是历史的必然。下面再略作具体分析：

（一）文化传统的影响 希尔伯特故乡的哥尼斯城堡建基于十三世纪，后来成为东普鲁士首都，那是一个著名的大学城。它位于布勒尔河两条支流之间，那里有桥联着一个岛和一个半岛，而数学史上那个著名的为欧拉解决的所谓“七桥问题”几乎是该城居民家喻户晓的一桩美谈。岛上有所古老的大学，还有大哲学家康德的墓地。显然，哥尼斯堡的自然环境和文化传统对于希尔伯特的成长来说是得天独厚的。

（二）家庭环境的影响 希尔伯特的父亲是一位普通的法官，母亲出生于普通商人家庭，但她爱好哲学、天文学和数学，特别对素数怀有浓厚的兴趣，这就影响了希尔伯特从小爱好数学。母亲每年在4月22日康德诞辰这一天，她总是带着小希尔伯特到康德墓地瞻仰康德的半身象，并且一字一句地拼读墙上刻着的康德的格言。这些对于希尔伯特从小爱科学，长大攀高峰，无疑会带来潜移默化的精神影响。

（三）社会舆论的影响 希尔伯特上小学二年级的时候，明可夫斯基一家从俄罗斯搬到了哥尼斯堡。明可夫斯基一家三兄弟当时称为三个“奇才”，以才能出众、性格迷人轰动了哥尼斯堡。特别是小神童明可夫斯基（Herman Minkowski 1864-1909）比希尔伯特小两岁，他的数学才能显著超过希尔伯特。他后来也成为大数学家，是数的几何（Geometry of Numbers）这一数论分支的创始人。

当时两家只有一河之隔。小明可夫斯基的数学才能出众对小希尔伯特不能不产生一种心理上的压力。确实，据希尔伯特后来写的回忆录来看，他承认自己小时候并非天才，而是一个较愚钝的孩子，当然数学才能远远在明可夫斯基之下。在希尔伯特的亲友中，也没有人提到过希尔伯特的能力曾受到过人们的注意。但是人们对明可夫斯基一家三兄弟的赞赏却激励了小希尔伯特。特别是小神童Hermann的数学天才象魔力一样征服了希尔伯特的心灵。

明可夫斯基刚满17岁时解决了“将正整数表成五个平方数和”的难题，同英国老数学家Henry Smith合得了法国巴黎科学院数学大奖，因而更加出名。当时希尔伯特的父亲还告诫希尔伯特，说不要同那样出名的人交朋友（以免被别人瞧不起）。可是希尔伯特不顾父亲的反对，毅然同明可夫斯基结成了终生最要好的朋友。

尽管希尔伯特和明可夫斯基在早年智力上有明显的差距，然而通过不断努力，希尔伯特后来不仅成为与明可夫斯基相提并论的大数学家，而且对整个数学的贡献还远远超过了明可夫斯基。这说明一个人的先天素质（所谓“天资”或“禀赋”）并不是决定成就大小的主要因素。先天素质的不足，可以在后天的实践中加以补偿。

（四）学校教育的影响 希尔伯特童年时代上的德国小学校，特别注意基础教育，非常强调语文、语法、算术等科目的基本

训练（尤其是语法一科，注重训练学生有条不紊地思维以及正确地表述思想的方式和方法）。希尔伯特进的中学和大学都充满自由学习的空气，这使他如鱼得水。

希尔伯特的青年时代是在哥尼斯堡大学度过的，那里有着浓厚的学术研究空气。著名的数学家Jacobi，Weierstrass，Weber，Lindemann等都在那里任教过，使该大学曾形成一个数学的研究中心。Lindemann曾以首先证明π为超越数而享有盛誉，他就是当年希尔伯特的学术导师。希尔伯特的学术论文原想研究"连分数的一种推广"。但经Lindemann指出，方知"那早已由Jacobi作出了"。在Lindemann的引导下，希尔伯特改搞"代数不变式理论"，结果大为成功。Lindemann对他的毕业论文极感满意。明可夫斯基在写给希尔伯特的信中也赞赏说："这样精彩的数学定理会出现在哥尼斯堡真是值得庆贺……"。可见，获得第一流的教师指导引路，也是希尔伯特成功的因素之一。

哥尼斯堡大学曾讲授一些最新颖的数学科目，这样就往往能把年轻人很快带到数学研究领域的前沿，从事创造性的工作。此外，启发式教学法对希尔伯特的教益也很大。例如，他曾选学线性微分方程课程，当时Fuchs教授的讲课方法与众不同。Fuchs习惯于在讲课时把自己置于危险困难境地(可能是缺乏备课习惯)，对要讲的内容总是现想现推。这样一来，就使得希尔伯特和他的同学们有机会瞧一瞧高明的数学思维过程是怎样进行的。

还有良师益友的互相切磋讨论，对希尔伯特的成长发展也起了十分重要的作用。当时希尔伯特和明可夫斯基的老师Hurwitz（他是一位只比希尔伯特大三岁的杰出数学家），非常器重两个学生。据说，每天下午准五点，三人必定相会，一起去苹果园散步，共同讨论问题，交流思想，交流研究心得。据希尔伯特后来回忆说，当年三个年轻人几乎考察了数学领域的每一个王国。可以想见，这应该是希尔伯特的才、学、识获得迅速成长的重要过程。假如没有这段经历，那末希尔伯特在1900年竟能在许多重要

领域中一次提出那样多的著名难题，倒是不易想象的了。

以上我们概述了希尔伯特成功的社会因素。当然，他本人的勤奋努力和艰苦奋斗等内在因素也是保证他获得成就的重要条件。例如，在研究工作中希尔伯特曾耐心地计算过四十多重的重积分。即使面对非常繁重的计算任务，他也是具有计算到底的坚强毅力的。事实上，很难设想缺乏坚强毅力的人能取得科学上的巨大成就（关于希尔伯特的详尽记载请参考C.Reid "Hilbert"一书）！

从方法论观点看，关于如何为青年人创造一个既有良师又有益友的环境，如何采用启发式方法讲授一系列新颖课程，并诱导青年人很快走上科研前沿等问题，显然能从希尔伯特成长的历史规律中，获得应有的解答。

§4　浅谈微观的数学方法论

每一个数学研究工作者都必须精通某些微观的数学方法论，才能有效地开展科研工作，获得丰硕成果。教师们也必须熟知这些方法论才能实行启发式教学法。

我们知道，美籍匈牙利数学家Pólya曾花数十年时间致力于"数学发现"与"解题思想方法"的研究。他的一些著作已被译成中文。特别值得重视的是他所著的《数学中的归纳与类比》（1954年出版）一书。在此书中作者曾选用不少富于启发性的例子说明归纳与类比方法如何成为发现数学真理的重要手段。

十八、十九世纪有突出贡献的数学家欧拉（Euler，1707-1783）和高斯（Gauss，1777-1855）都曾发表过一些经验之谈。欧拉说过："数学这门科学，需要观察，还需要实验"。高斯也提到过，他的许多定理都是靠归纳法发现的，证明只是补行的手续。

举例来说，欧拉关于多面体的面、顶、棱公式($F+V-E=2$)

显然就是从一批特殊的凸多面体的观察分析中归纳出来的。高斯青年时代曾著有《算术研究》（1801年出版的数论名著）一书，书中许多结果，包括著名的二次互反律等等，也都是首先从观察、实验、归纳过程中发现的。为什么数学真理如同物理科学领域中的定律和原理那样，有时可以通过实验与归纳方法去发现呢？原因很简单，因为数学对象本身(如数量关系与空间形式等）也具有客观实在性。

历史上许多有贡献的数学家，可以说无例外地都是善于应用“归纳法”与“类比法”去发现真理的能手。尤其是欧拉的许多发现与贡献早已经进入中学数学教材和大学低年级的课程之中，所以讲讲欧拉的一些光辉例子，对青年学生似乎更有教育意义。

为了说明类比法的作用，这里我们来介绍一下数学家伯努利(Bernoulli，1654-1705）的一个级数求和难题，是怎样被欧拉攻破的。伯努利是十七世纪杰出的数学家，他是古典概率论的创始人，对古典微积分学以及级数求和等问题都有贡献。但他没有办法算出自然数倒数平方的级数和$1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\frac{1}{25}+\cdots$。于是，他公开征求这一求和问题的解答，可惜直到他逝世时还未能见到有人解决此难题。这个难题过了数十年后才由欧拉解答出来。欧拉采用的方法就是一种巧妙的类比推理法。

首先，对于只含偶次项的$2n$次代数方程

$$b_0-b_1x^2+b_2x^4-\cdots+(-1)^nb_nx^{2n}=0\quad(b_0\neq0),$$

假设有$2n$个互不相同的根

$$\beta_1,\ -\beta_1,\ \beta_2,\ -\beta_2,\ \cdots,\ \beta_n,\ -\beta_n,$$

则得

$$b_0-b_1x^2+b_2x^4-\cdots+(-1)^nb_nx^{2n}$$

$$=b_0\left(1-\frac{x^2}{\beta_1^2}\right)\left(1-\frac{x^2}{\beta_2^2}\right)\cdots\left(1-\frac{x^2}{\beta_n^2}\right).$$

把乘积展开出来，易见x^2项的系数为

$$b_1 = b_0\left(\frac{1}{\beta_1^2} + \frac{1}{\beta_2^2} + \cdots + \frac{1}{\beta_n^2}\right),$$

以上所述都是代数方程式论中的初等知识。

欧拉考虑了三角方程

$$\frac{\sin x}{x} = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \frac{x^6}{7!} + \cdots = 0,$$

他把它看成是只含偶次项的无穷次代数方程。由于此方程含相异根$\pm\pi$，$\pm 2\pi$，$\pm 3\pi$，…，于是欧拉采用类比法，即仿照上述$2n$次多项式分解成乘积的形式，把这里出现的所谓无限次多项式也照样分解成因式乘积形式

$$\frac{\sin x}{x} = \left(1 - \frac{x^2}{\pi^2}\right)\left(1 - \frac{x^2}{4\pi^2}\right)\left(1 - \frac{x^2}{9\pi^2}\right)\cdots.$$

这便是著名的"欧拉乘积公式"。这样一来，再把右边的乘积展开出来便发现x^2项的系数是

$$-\frac{1}{3!} = -\frac{1}{\pi^2} + \frac{1}{4\pi^2} + \frac{1}{9\pi^2} + \cdots,$$

也就是

$$\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots = \frac{\pi^2}{6}.$$

这样，欧拉便完成了一项非常有趣的发现，给出了伯努利所未能找到的级数和。

据说欧拉发现上述结果后，当时并未能给出严格证明，不免有点又惊又疑。于是他做了数值计算验证，对等式两边分别进行计算，得出的数值都等于1.644934…。算到第七位数字都一致，这才使得欧拉确信他的发现正确无疑。（这和实验物理学家发现物理定律时的态度多么相似！）当然，现今的数学分析教程中已有各种方法可以证明上述结果。

顺便提几句，在现代初、高等数学教育中，特别反映在教材与教学方法中，似乎过于偏重演绎论证的训练，把学生的注意力都吸引到形式论证（逻辑推理）的"严密性"上去，这对于培养学

生的创造力来说实际是不利的。当然，必要的逻辑推理训练不可少；但对于有作为的数学工作者来说，发现和创新比命题论证更重要。因为一旦抓到真理之后，补行证明往往只是时间问题。大数学家高斯早就谈过这种经验。当然也有例外，例如数学上有许多诱惑人心的“猜想”，看来似乎是“真”的，但却证明不了。其实很多猜想未必真正抓到真理，所以事后被证明是错的也不少。

归纳法与类比法是数学方法论中最基本的方法之一，用好了能获得新的成果，乃至完成重要发现。但要真正用好也不容易。首先，要有敏锐的观察力，才能从众多的特例中归纳总结出一般性命题来。“特例”有时是现成的，有时却需要故意构造出来。要用好类比法需要有较丰富的数学知识，知识面越广，在数学思维中可用作类比推理的题材就越多，因而能形成普遍命题的机会（或发现数学一般真理的机会）也就越多。很难设想，知识面很窄的人能完成重大的发现。事实上，利用类比法形成普遍命题的过程是通过“联想—预见”来完成的。联想就要靠已有的知识为基础。

一般说来，归纳与类比在从事数学创造性科学研究活动过程中的作用如下图所示。

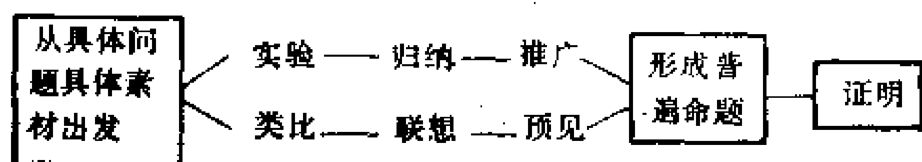


图1-1

形成的“普遍命题”在完成证明之前往往是一种猜想，因此，只有经过严格证明之后才能成为确定的定理或论断。但在许多情况下，“推广”和“预见”的过程中已经蕴含有普遍命题的直观论证或不甚严格的证明。这样，形成普遍命题的这一重要步骤实际已经完成了数学真理的发现工作。当然，从发现到证明有时也往往需要走一段艰苦的路程。

数学史上许多杰出的数学家往往既是发现与发明的能手，又是精于证明技巧的硬手。但是也能遇见这样两种数学家：一种是擅长于数学发现的专家，另一种是专长于论证的专家。就创造性数学思维来说，前者擅长于“发散思维”，后者较精于“收敛思维”。

在数学的创造性工作中，“抽象分析法”也是一种常用的重要方法。例如，欧拉解决哥尼斯堡七桥问题时，就是采用了这种方法。欧拉的解法在许多书上都有介绍。解决的基本步骤无非是：把人们步行过桥的问题经过分析，抽象成为一个“一笔画”的问题，即一笔能否画出如下图形的问题。

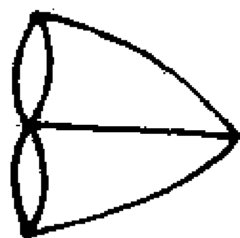


图1-2

欧拉原来是这样想的：既然岛与半岛无非是桥梁的连接地点，两岸陆地也是桥梁通往的地点，那末就不妨把四处地点缩小（抽象）成四个点，并把七条桥表示（抽象）成七条线，这样当然并不改变问题的实质。于是，人们企图一次无重复地走过七条桥的问题即等价于一笔画出上述图形的问题。这样的分析思考方法，就叫做“抽象分析法”或“数学模型法”。这里，一笔画问题中的几何图形就是七桥问题的数学模型。

接着，欧拉又考察了一笔画的结构特征，一笔画有个起点和终点（特别，起点与终点重合时便成为自封图形）。除起点与终点外，一笔画中出现的交点处曲线总是一进一出，故通过交点的曲线总是偶数条。如此说来，一笔画中至多只有两个点（即起点与终点）有可能通过奇数条曲线。我们看图1-2，立即发现四个点都通过奇数条曲线。因此，可以断言它不是一笔能够画出的图形。

抽象分析法还能用来确立新的基本概念，导致数学新学科或新分支的产生。大家知道，随着现代计算机的发展已经产生了一门新的数学分科，叫作“计算理论”。英国数学家图灵（A.M. Turing，1912-1954）的工作在这整个历史发展过程中起着关

键性的作用。要不是图灵当初彻底分析了计算的实质，并从理论上论证了“通用计算机”的可能性，也许当年电子计算机的发明与发展不会那样顺利而迅速。

什么是计算？多少世纪以来，人们都学习计算，经常使用计算，但在1936年前，从未有人对“计算”的本质进行过深刻的分析。图灵就是应用抽象分析法首先阐明计算本质的一位数学家。

我们仔细地观察不难发现，一个人进行笔算时总是把一些符号写在纸上，当计算中出现不同的特殊符号时，就改变作计算的动作。而计算者工作时用的是铅笔还是钢笔，用的纸是有行的、无行的或方格纸等，这些都与计算过程的实质无关。图灵在分析计算过程时，正是对过程中的一切无关因素加以舍弃，对过程进行去伪存真、去粗取精，才发现了计算的本质。这样才导致后来通用电子计算机的发明。

图灵分析了人们在从事计算时所遵循的最基本的法则。首先，他发现计算者可以把一切计算内容写在一条线性的带子上，至于通常使用的纸的二维性质并不是本质的。如果你愿意，可以不用普通纸带，而改用录音机用的那种磁带（当然，在后一情况里出现的符号将是磁信号而不是纸上写的符号，但在概念上并无本质区别）。所以，图灵由抽象分析获得的一个重要结论是：一切实际计算过程都具有“线性”性质。

计算过程的本质既是线性的，就不妨假定线性的带子上划分为若干方格。我们知道，一切有理数均可采用二进位数表出。所以在图灵理想的计算机上，不妨规定线性带子的每一方格内可以记上一个符号0或1。

图灵对计算过程所作的第二步抽象分析是，一切计算过程的实质无非是每一步把在方格里看到的0换成1，或者把1换成0，或者有时需要把注意力转移到另一方格上去，不妨假定注意力的转移只限于从所看到的方格移到左右相邻的方格（这对计算过程的实质并无真正限制）。

经过如上的抽象分析后，图灵便得出这样的结论：任何计算都可以看作是由一个人工计算者（或计算机器）来做的，它使用线性带子上成串的0和1，不外乎执行下列各指令：（1）写符号0；（2）写符号1；（3）向左移一格；（4）向右移一格；（5）观察现在扫描的符号并相应选择下一步骤；（6）停止。计算者所执行的程序，也就是这类指令所排列成的形式表。这样分析之后，计算过程的实质也就彻底搞清楚了。

近代应用数学的另一重要分支是“信息论”。什么是信息？信息的含义是极其宽广的。不用说是人类，即使是一些禽兽间也以它们的特殊方式（如鸣叫声）传递信息。在现代人类生活中，可以说处处在传递着信息，无论是报章、杂志、电视、电报、电话、广播乃至音乐演奏等都包含着信息的内容。所有这些，都有一个共性，即信息总是可以传递而且必须传递的。

信息既然要传递，就得考虑如何衡量信息量。这就需要建立信息的度量理论。怎样界定信息的度量概念呢？这又必须应用抽象分析方法。下面针对较直观的例子来分析。

以英文的26个字母为例，试问如何衡量它的信息量？如果只允许一个字母A出现，其它字母均不许出现，则信息量不妨规定为1，即定义为一个单位。在各个字母以等概率出现的情况下，26个英文字母所包含的信息量可以认为是26。又如果考虑由两个字母所拼成的字（假定每种拼法都赋予意义），则字的集合共含26×26个元素。根据信息与集合元素个数成正比的原则，该集合的信息量应该是26×26。但信息量的这种按“乘法增加”，显然与直观预期不协调。直观上总希望按“加法增加”。因而可采用对数来衡量，即规定26个字母的信息为$\log 26$。从而上述字的集合的信息量便为

$$\log(26\times 26)=\log 26+\log 26.$$

现在剩下的问题是，这个对数应以什么数为底？

正如许国志先生在《略谈应用数学的范畴》一文中所说的，中

国古代的烽火台是最简单的信息传递的例子。“无火报平安，有火敌来袭。”这仅包含两个信息。因此，如果对数以2为底，那么烽火台所提供的信息便是$\log_2 2=1$。由于任何事物的肯定与否定以及由于语言表述与逻辑的二值性都只包含两个信息，所以用2作底应该是最为合理的。于是26个字母的平均信息量便是$\log_2 26=4.7$。假定各字母出现的概率相同，则每个字母出现的概率都等于1/26.因此上述信息量也可记成：$-\log_2(1/26)=4.7$。然而一般情形下，26个字母出现的概率可以各不相同.如果以p_1，p_2，…，p_{26}分别代表字母A，B，…，Z出现的概率，那么按下述加权平均

$$-(p_1\log_2 p_1+p_2\log_2 p_2+\cdots+p_{26}\log_2 p_{26}),$$

算出的数值即等于诸字母的“平均信息量”。特别，诸字母以等概率1/26出现时，则由上式仍可得出$-\log_2(1/26)=4.7$。我们知道，上述公式正是信息度量理论中的基本公式（这个公式推广到一般集合的情形当然是显而易见的事）。

在上述分析过程中，我们事实上用到了从特殊到一般、从具体到抽象的思想方法。

近代数学中比较普遍采用的“公理化方法”和应用数学领域里经常使用的“模型方法”，其实也都是抽象分析法的具体运用和表现。这些需要另辟专题，在后面将详细论述。

最后，我们对有兴趣钻研数学方法论的青年数学工作者，特提出如下几点希望和建议以供参考：（1）最好能抽出些时间主动阅读一点数学发展史，以加深对数学发展宏观规律的认识。（2）尽可能选读一些著名经典作家（数学家）的全集或选集中的若干代表性作品，以便领会某些卓越的心智活动法则和规律。（3）在可能范围内，最好能在数学科学（甚至是自然科学）的广阔领域中博览群书，以开拓自已的知识疆域，俾有利于发展自已的理解能力和想象能力。（4）宜通过辩证法的学习，尽早确立科学的反映论观点。

第2讲 略论数学模型方法

§1 数学模型的意义

"数学模型方法"（mathematical modelling method）简称MM方法，它不仅是处理数学理论问题的一种经典方法，而且也是处理科技领域中各种实际问题的一般数学方法。特别，现代电子计算机的广泛应用和科学技术的数学化趋势，使得MM方法已经非常广泛地应用于自然科学、工程技术科学与社会科学的一切领域中。例如，经济科学、军事科学、交通运输等管理科学领域，都无例外地应用着MM方法。

如所知，现代各门应用数学所以具有解决实际问题的功能，主要就是通过提供数学模型方法而显示出来的。所以，凡从事应用数学的科研工作者，都必须精通MM方法。

"数学模型"的含义很广。粗略说来，数学模型乃是针对或参照某种事物系统的特征或数量相依关系，采用形式化数学语言，概括地或近似地表述出来的一种数学结构。当然，这种数学结构应该是借助于数学概念和符号刻划出来的某种系统的纯关系结构。所谓纯关系结构是指已经扬弃了一切与关系无本质联系的属性后的系统而言，所以，在数学模型的形成过程中，已经用了抽象分析法。也可以说，抽象分析法是构造数学模型的基本手段。

仔细说来，数学模型有广义的解释和狭义的解释。从广义上讲，数学中各种基本概念，如实数、向量、集合、群、环、域、范畴、线性空间、拓扑空间等等都可叫作MM，因为它们都是以各自相应的现实原型（实体）作为背景而加以抽象出来的最基本的数学概念。这些可称为原始的MM。

例一 欧氏几何是关于直觉空间形体（刚体运动下图形结构不变的形体）关系分析的MM。

例二 自然数1，2，3，…，n，…是用以描述离散数量的MM。

例三 每一个代数方程式或数学公式也都是一个MM。例如，$ax^2+bx+c=0$就是一类具体应用问题的MM。

总之，按广义的解释，凡一切数学概念、数学理论体系、各种数学公式、各种方程式（代数方程，函数方程、微分方程、差分方程，积分方程……）以及由公式系列构成的算法系统等等都可称之为MM。

但按狭义的解释，只有那些反映特定问题或特定的具体事物系统的数学关系结构才叫作MM。例如，在应用数学中，MM一词通常都作狭义解释，而构造MM的目的就是为了解决具体实际问题。

§2 数学模型的类别及简单例子

如前所述，在现代应用数学中，数学模型往往是对特定对象系统中的量的关系方面的模写。由于特定问题是形形色色、千差万别的，因此，针对具体问题具体对象去建立MM时，必须进行具体分析。大致说来，MM可分三大类。

一类是**确定性数学模型**。这类模型所相应的实体对象（又称背景对象）具有确定性或固定性，对象间又具有必然的关系。这类模型的表示形式可以是各种各样的方程式、关系式、逻辑关系式、网络图等等。所使用的方法无非是经典数学方法。

二类是**随机性数学模型**。这类模型的背景对象具有或然性或随机性。MM的表示工具无非是概率论、过程论及数理统计学等方法。

三类是**模糊性数学模型**。这类模型所对应的实体对象及其

关系均具有模糊性。MM的基本表示工具便是Fuzzy子集合理论及Fuzzy逻辑等等。

当然，从复杂多变的实际对象关系中分析结晶出来的数学模型也可能是兼具随机性和模糊性的混合型数学模型，表现这类MM所使用的数学工具也就不能是单纯的一门数学了。

为了便于说明MM方法，下面举几个简单且熟知的例子，目的无非是为了总结出一般性的思想方法。

例一 哥尼斯堡七桥问题。

欧拉解决这个问题的思想方法，其实也就是MM方法。七桥问题这样一个实际问题，就是MM的现实原型。通过抽象分析，七桥问题被抽象成线路拓扑的一笔画问题。后者便是前者的M M。建立MM之后，便可在MM上进行逻辑推理和论证，推出的结论是无解。然后再把这个结论对应地翻译回去，即返回到现实原型上，便得到实际问题无解的答案。

上述思想方法可用框图表示如下：

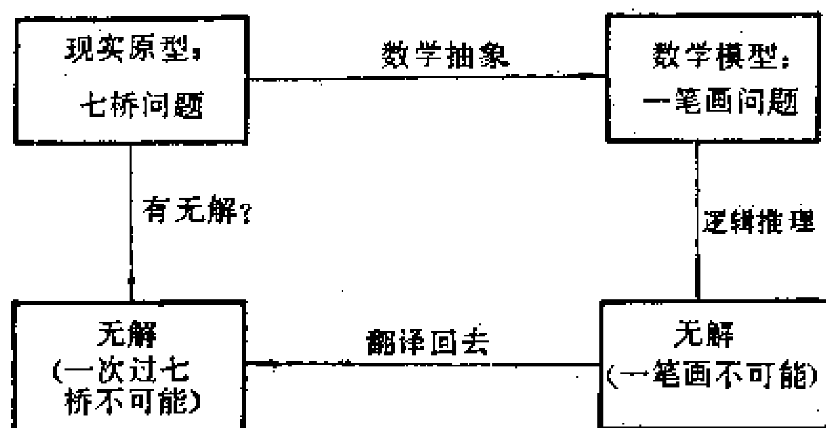


图2-1

显然，这个例子中的MM是个确定性模型。

例二 关于物体冷却过程的一个物理问题：设某物体置于气温为24°C的空气中，在时刻$t=0$时，物体温度为$u_0=150°C$。经过

10分钟后物体温度变为$u_1 = 100°C$。试决定该物体温度u与时间 t 之间的关系并计算$t = 20$分钟时物体的温度。

为解决此问题就需要构造一个MM。首先，要对问题作定 性分析。鉴于该问题仅涉及必然性现象，故应作确定性数学模型。又为了能用形式化数学语言表述MM，而这个MM反映的是物理现象（物体冷却过程），故还必须应用物理学定律——牛顿冷却定律：热量总是从温度高的物体向温度低的物体传导的。在一定温度范围内，一个物体的温度变化速率恒与该物体和所在介质之温差成正比。

解此问题的思路类似于例一，即如下图所示。

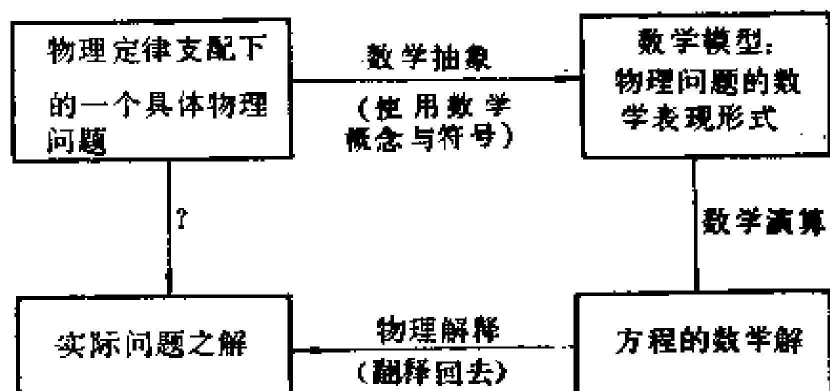


图2-2

在上述问题里，物体温度u应是时间变量的连续函数，不 妨 记为$u = u(t)$。对初始温度u_0而言，温差为$u_0 - u_a$（u_a为空气介 质 温度）。我们又知道，应变量（函数）的变化率可用微商（导数）概念来表述。于是，物体冷却过程（现实原型）的数学模型就是如下形式的微分方程：

$$\frac{du}{dt} = -k(u - u_a),$$

这里k是比例常数，在具体的问题里可以确定下来。

具体问题要求找出函数关系$u = u(t)$的显式表示。为此，解上述微分方程，易得

$$\int \frac{d(u-u_a)}{u-u_a} = -k\int dt = -kt+c,$$

$$\log(u-u_a) = -kt+c,$$

$$u-u_a = Ae^{-kt}, \quad A\text{——常数。}$$

按初始条件$t=0$：$u=u(0)=u_0$，故得

$$u_0-u_a = Ae^0 = A,$$

也即有　$u-u_a=(u_0-u_a)e^{-kt}$

或　　　$u=(u_0-u_a)e^{-kt}+u_a,$

这便是所要寻求的方程解，它是冷却过程数学模型的显式表示。

有了上述一般性模型之后，只须把实际问题里的具体数据一一代入，即可得出

$$100=(150-24)e^{-10k}+24.$$

由此确定出$k\doteq0.051$。因此针对具体问题的特殊模型为

$$u=24+126e^{-0.051t}$$

特别，命 $t=20$代入则得

$$u(20)\doteq24^\circ+40^\circ=64^\circ.$$

这便是所要寻求的问题答案（这里我们对方程中出现的数量已翻译为物理量，例如64°即表示摄氏64度）。

例三　布丰（Buffon）投针实验的数学模型问题（这个著名例子在现代概率论教程里经常谈到）：试在平地上划出一组间隔距离为一寸的平行线。以一寸长的针（质量均匀的细针）随机地掷到画有平行线条的地面上，问针与平行线接触的机会有多大？

首先分析一下：现实原型（实体对象）具有随机性、或然性。故所求MM必为一随机模型（或概率模型）。如图示，$0\leqslant l\leqslant1$，$0\leqslant\theta\leqslant\pi$，截长$l$与交角$\theta$

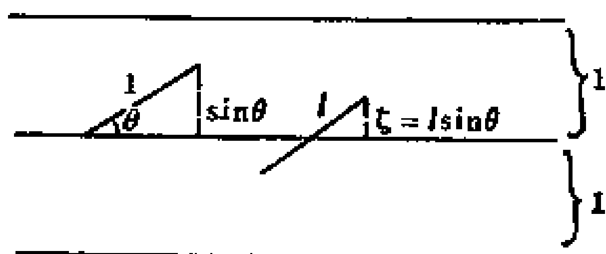


图2-3

显然都具有随机性，故l、θ均为随机变量，且各取可能值的机会

均具有“等或然性”。

在所讨论的问题里，现实原型为投针实验问题，即细针触线的机会问题。通过引进随机变量θ与$\xi=l\sin\theta$也就可表述为一个概率计算问题，后者便是前者的数学模型。以P表概率，E表概率均值（数学期望），则相应的MM即可表述为

$$\underset{\theta}{E}\left\{P\Big(\xi\in(0,\sin\theta)\Big)\,\Big|\,0\leqslant\theta\leqslant\pi\right\}=?$$

经过简单计算可得

$$P(\xi\in(0,\sin\theta))=\sin\theta,$$

$$\underset{\theta}{E}\left\{\sin\theta\right\}=\frac{1}{\pi}\int_0^\pi\sin\theta d\theta=\frac{1}{\pi}\Big[-\cos\theta\Big]_0^\pi=\frac{2}{\pi}.$$

因此投针问题所要寻求的机会（概率）即等于$2/\pi$。

我们知道布丰曾利用实验验证了上述结论（即由MM导出的结论）的实际可靠性。可见数学理论分析出来的模型确实能反映客观规律性。

§3　MM的构造过程及特点

从前面两节我们已经知道，数学模型与现实原型的关系是反映与被反映的关系。因为MM的构造须经过抽象分析过程，即经过对现实原型扬弃次要环节的过程，故MM和现实原型只能具有相对一致性，事实上往往只能在基本环节上（这些环节或关系结构是能借助数学形式表现出来的）有着近似的一致性。就这一点来说，MM有点象美术中的写生画或文学中的人物典型塑造那样，都无非是某些现实原型的本质或特征的近似刻划与集中反映而已。

一般说来，构造MM的基本过程可分如下几个步骤：

第一步：对所研究的实际问题即现实原型，要分析其对象与关系结构（包括量变因果关系）的本质属性，以便确定其MM的

类别。

第二步：要确定所研究的系统并抓住主要矛盾。为使用MM方法，须考察问题所属系统，如热力系统、水力系统、电力系统、生态系统、生产管理系统、市场供销系统、人体神经系统等等。在确定系统的过程中有时还须从大系统中分离出子系统来。例如，研究潮汐摩擦问题时，须确定地-月系统。这个系统须从太阳系（大系统）中分离出来。

为建立MM还必须选择具有关键性作用的变量或量的关系进行考察。这就是说，要抓住主要因素，因为MM所反映的应该是主要因素间的关系结构。

第三步：要进行数学抽象。对事物对象及诸对象间的关系都要进行抽象，要尽可能使用数学概念、数学符号和数学表达式去表现事物对象及关系。如果遇到现成的数学工具不够用时，则还需要大胆创造，根据实际情况，提出新的数学概念和数学方法去表现MM。例如，六十年代中期L.Zadeh创始的“Fuzzy数学”就是为了处理模糊事物对象的MM问题所创始的一门数学。

上述第二、第三步是构造MM的基本步骤。

例一 试考虑地球运动问题。首先，应确定研究的系统是日-地系统，它是宇宙力学系统的子系统。在这系统中具有关键作用的各种量的概念有：质量、位置、距离、时间、速度、加速度等等。此外，还有运动轨道也是必须考察的重要因素。当然，所有这些对象或因素都可借助于数学上的常量、变量、坐标、微商、曲线等加以表述。在这个系统中，要抓住的主要矛盾就是太阳与地球的相互作用力。因为相对说来，其它诸星体对地球的引力作用毕竟是微不足道的次要因素，故可忽略不计。

其次，太阳是恒星，其位置可以当作不变。日、地都可看作是几何上的点。这些便是为了作出MM必须采取的数学抽象方法。

最后，为了作出地球运动的MM（即运动方程式）还必须利

用普遍的自然科学定律，即万有引力定律及牛顿的第二运动定律。试以$\vec{r}$表示地球点的坐标向量（即位置向量），这是一个依赖于时间t的矢变量。又以r表$\vec{r}$的长度。这样，便得如下方程式，即MM：

$$\frac{d^2}{dt^2}\vec{r} = -\frac{1}{r^2}\left(\frac{\vec{r}}{r}\right)K(K\text{——万有引力常数})。$$

一般说来，构成数学模型之后，还必须把从MM上分析得出的理论解答（数学推论）返回到现实中去，看看能否确实回答实际问题。当然，当利用MM去解答实际问题时，还必须把实际问题里所涉及的若干特殊条件加入进去一并考察。比方，假如建立的MM是一个微分方程，则必须把某个实际问题中提供的初始条件或边界条件一并考虑，这样才能获得合乎某个实际问题需要的解答。

按照上述三步骤构造出来的MM至少应具有两个特点：（1）在MM上应具有严格推导（逻辑推理）的可能性以及导出结论的确定性。假如没有这一特点，则MM就没有什么用处了。（2）MM相对于较复杂的现实原型说来，应该具有化繁为简、化难为易的特点。因此在用以解决实际问题时，应该容许有一定的误差范围。好的MM还应该具有估计误差范围的性能。

判断一个数学模型是否有效或是否合用，主要应该衡量上述两个特点。如果对同一问题的两个MM具有同样效能，则比较简单易用的MM便是较好的MM。

§4　怎样培训构造MM的能力

从数学模型的性质要求和构造方法来看，一个人构造MM的能力至少包括四个方面：一是理解实际问题的能力，二是抽象分析的能力，三是运用数学工具的能力（包括运用数学形式语言的能力），四是通过实践加以验证的能力。所以，为了培训构造

MM的工作能力，首先必须学习有关的自然科学、工程科学或社会科学的某些分支领域的知识，要掌握好这些领域中的定律、法则和规律，这样才能有助于提高构造有关领域中的MM的实际工作能力。其次，在学习各门数学时，要注意多做些应用题（例如，在中学阶段应多做些代数方程应用题，在大学阶段应多做些微积分与微分方程方面的应用题，在力学课程中也应多做应用题）。这对提高分析能力和运用数学语言能力是必不可少的基本训练。最后，还要多多接触实际问题。有时还需要深入到实际具体部门中去，才能培养善于洞察问题关键的能力。

由于现代科技各部门的分工越来越细，而解决大型实际问题的MM往往需要几个不同部门知识与技术的高度结合，于是跨行业专家间的协作就成为必要的了。但是，这种协作的有效性程度又往往取决于共同理解问题性质的平均深度。因此，不论哪一行都要成为构造MM的能手或有效合作者，除了精通本行专业之外，还必须要有较宽广的科技知识修养。这一点，对培养应用数学工作者说来似乎是特别需要注意的。例如，现代西德有些大学研究所培养运筹学科的硕士生时，就已经充分注意到这一点。

第3讲 关系映射反演原则的应用

§1 何谓“关系映射反演原则”？

所谓**关系映射反演原则**，是指一种分析处理问题的普遍方法或准则。它是属于一般科学方法论性质范畴的一种工作原则。因为这种普遍方法或工作原则包括着对所要研究的问题中的关系结构，采取映射和反演两个步骤去解决问题，所以给它的命名也就不免有点冗长了。这里所说的映射和反演可以赋予很广泛的含义。

在第2讲中讨论过的数学模型方法，也就是利用形式化的数学模型去反映（模写、刻划、表征）现实系统中的关系结构，然后利用通过对模型的逻辑分析演绎得出的结论，把它反演（翻译）回去解答现实原型中的某些问题。这种处理问题的过程已经体现了关系—映射—反演的思想方法。因为从给定的系统（现实原型）到数学模型的某种对应关系，即可理解为一种映射关系。把分析模型得出的理论分析结果（或结论）又对应到现实原型上去，从而给出实际问题的解答，这就是一个反演过程。

其实，即使在日常生活中，人们也经常自觉或不自觉地在运用着关系映射反演原则。比如说，一个人对着镜子剃胡子，镜子里照出他脸颊上胡子的映象，从胡子到映象的关系就叫作映射。所以，映射就是连系着原象和映象的一种对应关系。他用剃刀修剪胡子时，作为原象的胡子和剃刀两者的关系可以叫作原象关系。这种原象关系在镜子里表现为映象关系。他从镜子里看到这种映象关系后，便能调整剃刀的映象与胡子的映象的关系，于是，他也就真正修剪了胡子。这里显然用到了反演原则；因为他

已经根据镜子里的映象能对应地反演为原象的这一原理，使剃刀准确地修剪了真实的胡子（原象）。

为了缩短名词的称呼，我们不妨把关系(relationship)映射(mapping)反演（inversion）原则简称为RMI原则。在此先对这一原则概略说明如下：令R表示一组原象的关系结构（或原象系统），其中包含着待确定的原象x。令M表示一种映射(一一对应法则)，通过它的作用假定原象结构系统R被映成映象关系结构R^*，其中自然包含着未知原象x的映象x^*。如果有办法把x^*确定下来，则通过反演即逆映射 $I=M^{-1}$也就相应地把x确定下来。这便是RMI工作原则的基本内容，可用框图表示如下。

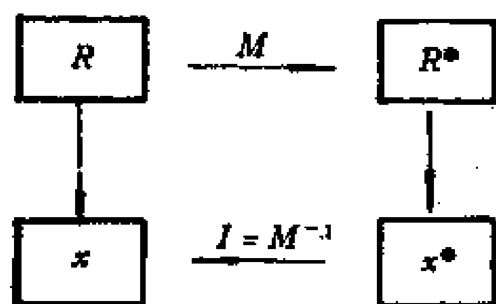


图3-1

我们知道，不仅在数学中，而且几乎在一切工程技术或应用科学部门中，都往往利用这一原则去解决问题。通常总是选择最合适的映射M，使得待定原象x（即具体问题中的目标对象）的映象x^*较容易地确定下来，从而通过反演也就较容易地把目标对象x寻找出来。由于许多问题里的x往往是不容易确定的，因此上述工作原则也就非常有用。

我们说过，映射和反演可以赋予很广泛的含义，因此RMI原则实际可以理解为一种包罗万象的科学方法论原则。例如，概念乃是人脑活动的最高产物，它是事物对象或事物关系在人脑中的反映。我们不妨把概念的形成过程看成是通过人脑机制活动完成的映射。于是概念便是事物原象（对象及关系）的映象。利用概念思维（包括逻辑分析推理）得出的结论，返回到事物原型上去解决问题，这可以理解为一种反演过程。这样看来，一般情况下通过人脑概念思维去解决实际问题的全过程，在一定条件下也可以看作是RMI原则的应用。

在上述例子中，R，R^*，M，…等可赋予下述含义：

R——包含着实际问题的事物关系系统；

R^*——包含着理论问题的概念关系系统；

M——从事物关系到概念关系的形成过程；

$I=M^{-1}$——从概念返回实际事物的逆过程；

x^*——实际问题中的未知目标x的映象。

这里运用RMI原则的目的是要把实际问题中的某个待求的结论确定下来。

以上所论是假定未知结论（或未知原象）x可以从R关系结构中确定下来。但直接确定x较为困难，因此不得不拐一个弯，通过R^*找出x^*后再由x^*反转到x。当然，这里用到的映射M与反演I必须是切实可行的，否则仍无济于事。

但是，人们通过概念思维（概念逻辑演绎思维）去解决某些实际问题的过程，有时要比上述情况复杂困难得多。例如，包含着实际问题x的某个事物系统R可能不是一个**静态的关系结构**，而是一个发展着的需要继续补充条件的**动态关系结构**，而且只有增补某些条件后才能把x的结论确定下来。这时，R可以看成是处于初始状态的关系结构，R^*是这种结构的映象。但R^*尚不足以确定x的映象x^*。为了确定x^*需要假定在R^*上补充一组条件C^*。相应地，也就需要在R上追补一组条件C。这样，用以解决某个实际问题的RMI原则的运用过程就可用如下的框图来表示。

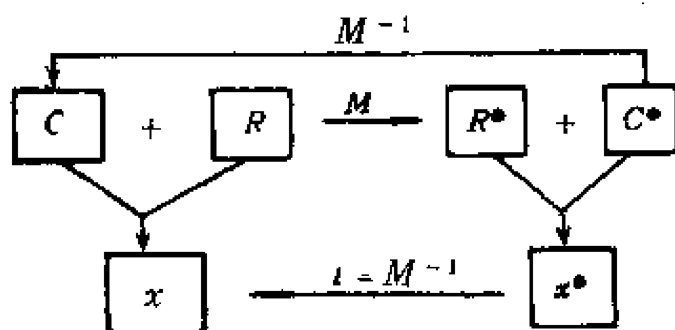


图3-2

我们从事数学研究工作时，特别是企图解决实际中提出的应

用数学问题时，为了得到待求的答案x，往往采用如上形式的RMI方法原则。因为在原来给定的关系结构系统R里往往由于关系不够丰富（即条件不够充分），尚不足以确定x；因此在工作过程中需要通过逆推法，求出一组条件C^*之后，再相应地把条件C追补到R上，然后保证由x^*反演成x。结论便是：由R添加C可以导出预期的x。

上述的思想方法原则，对于一个面临战斗任务的军事指挥员来说也非常有用。可以说，历史上每一个善于取胜的军事统帅都是自觉或不自觉地善于运用上述方法原则的人。法国历史上的拿破仑便是一个典型例子。例如在著名的Austeritz战役中，他曾统率法军打败兵力占绝对优势的俄奥联军的联合进攻，并使联军全面崩溃。这是军事史上恩格斯特别称道的一个例子。

拿破仑当年的思想方法或可表述为如下的模式：首先是两军当时的布阵形势加上地形、季节等各种自然因素便组成动态的关系结构R。这个R在拿破仑头脑中的映象为R^*，它是一系列形象、概念思维的产物。当然R^*并不足以导致想象中的胜利目标x^*。正是拿破仑卓越的洞察力（甚至把敌方统帅的心理状态也作为因素考虑在内）预见到敌方必将采取的军事行动路线，并相应地采取了击溃和消灭敌方的行动时间计划。所有这些想法和计划便组成补充条件C^*。于是R^*加上C^*便可保证从逻辑推理上导致胜利目标x^*。由于这些概念推理和行动计划都能在实践中体现出来，所以果然取得了预期的胜利（所期待的实际答案）。

作为属于一般方法论范畴的RMI原则的讨论到此为止。往下我们将专门讨论数学领域中的RMI原则的表述形式及其应用方面的一系列例子。

§2 数学中的RMI原则

因为数学是一门精确的科学，所以只要一进入数学领域，不

管是任何方法或原则均可获得比较确切的表述形式。下面我们先用描述方式给出一系列常用名词的解释。

数学对象 凡可表述为数学概念的事物(对象)个体称之为数学对象。例如，数、量、数列、向量、变数、函数、方程、泛函、函数族、点、线、面、几何图形、空间、集合、运算、算子、映射、随机变数、概率、分布、测度、级数、导数、积分、模糊集、群、环、域、范畴、代数系统、基数、序数、邻域、单子、非标准实数、数学模型、滤集等等。

一般说来，数学对象须具有三个特点，一是**一意确定性**，二是**逻辑演绎性**，三是特殊情况下的**客体背景存在性**。可见，凭空虚构的概念不能成为数学对象。

关系结构 (relation-structure) 由一些数学对象构成的集合称之为无关系结构。如果在集合的元素（对象）间存在着某种或某些数学关系则称为关系结构（所谓**数学关系**是指在数学对象间可以确切定义的关系。特别，例如有代数关系、序关系、拓扑关系、函数关系、泛函关系、相容关系、不相容关系等）。

一般说来，数学关系结构具有下列三条件：一是结构系统中的对象必须是数学对象，二是对象间的联系必须是数学关系，三是结构系统具有某种整体性或可分解性。

映射（或**变换**） 凡是在两类数学对象或两个数学集合的元素之间建立了一种“对应关系”，就定义了一个映射。特别，如果是一一对应关系，则称为**可逆映射**。例如，代数中的线性变换，几何中的射影变换，分析学中的变数代换、函数变换、数列变换、积分变换，以及拓扑学中的拓扑变换等都是映射概念的熟知例子。

假设φ是一个映射，它把集合$S=\{a\}$中的元素映入(或映满)另一集合$S^*=\{a^*\}$，其中a^*表示a的映象，a称为原象，这时可记作

$$\varphi: S\longmapsto S^*, \ \varphi(a)=a^*.$$

如果S还是一个关系结构，φ能将S映满S^*，则可记$S^* = \varphi(S)$，并称S^*为映象关系结构。

特别，假如关系结构S中包含一个未知性状的对象x，它是我们问题中需要确定其性状的目标，则称x为**目标原象**，在映射φ作用下，$x^* = \varphi(x)$ 便称为**目标映象**。

如果目标映象能通过确定的数学方法从映象关系结构系统S^*中确定出来（或寻找出来），则称映射方法φ为**可定映映射**。于是，数学中的RMI原则便可陈述如下：

给定一个含有目标原象x的关系结构系统S。如果能找到一个可定映映射φ，将S映入或映满S^*，则可从S^*通过一定的数学方法把目标映象$x^* = \varphi(x)$确定出来，从而通过反演即逆映射φ^{-1}便可把$x = \varphi^{-1}(x^*)$确定出来。这个过程的框图如下所示。

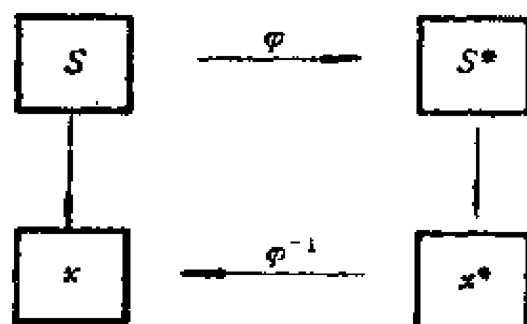


图3-3

全过程包括的步骤为：关系—映射—定映—反演—得解。

§3　若干较简单的例子

正因为RMI原则是一个极普遍的方法原则，所以无论在初等数学中或是现代高等数学领域中都可以找到它的许多应用实例。为节省篇幅计，我们只介绍一些极典型的例子。

例一　人们进行庞大数字的开方、乘方等数值计算时，往往应用对数方法。比方，要计算 $p = a^{\frac{1}{3}} b^{\frac{1}{7}}$ 的数值，计算过程即如下图所示。

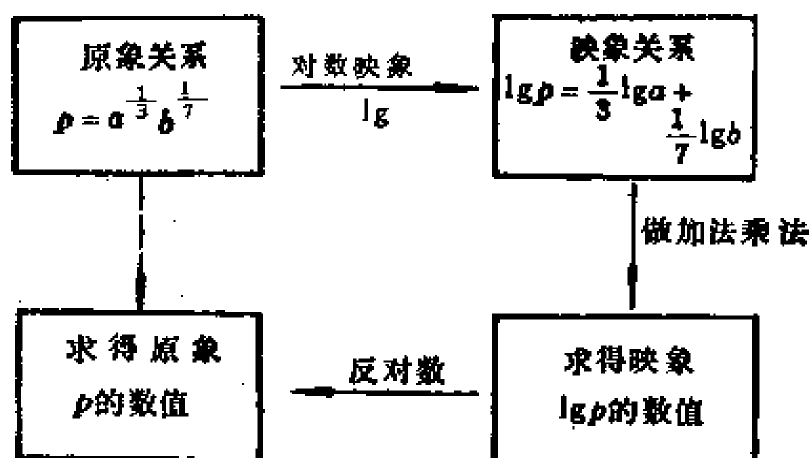


图3-4

在这个例子中，映射的运用是别开生面的：不是直接求p值，而是先求p在某种映射下的映象。显然，这里映射方法的选取非常成功。纳皮尔（Napier）的贡献就在于他看透了指数运算与真数运算的对应法则（映射与反演的关系），把后者的计算任务转化为前者的计算任务（把乘法转化为加法，把开方根计算转化为除法或分数乘法），从而大大提高了计算效率。纳皮尔在十六世纪末期首先把上述映射及其反演发展成一套数值计算方法，并编制了对数表作为映射反演的工具。这就是他在数学史上不可磨灭的功绩。

例二 利用射影变换的不变性，可以将平面几何中的许多定理推广到射影几何中去。今以著名的巴斯卡（Pascal）定理为例来说明其思想方法。这个定理断言："圆内接任意六边形的三双对边延长线的交点必共线"。把圆换成椭圆，这定理照样成立。这就是射影几何中的巴斯卡定理。

给定一个椭圆和它的内接六边形。于是三双对边的延长线给出三个交点A，B，C（假定其中没有无穷远点）。今在它们两两之间用直线连结起来，可得到一个未知性状的几何物，即未知原象（所谓"未知"是指该几何物是一三角形还是一直线尚未确定）。上述整个几何图形所表现的关系结构便是原象关系。为了确定未知原象的几何性状，可对上述原象关系实行射影变换 即

把上述椭圆置于一个适当的圆锥的截面上，使之正好成为截曰。这时圆锥底面上的对应图形（射影）是一个圆内接六边形。由于平面几何中的巴斯卡定理已知六边形三双对边延长线的三个交点共线，于是根据射影变换下共线关系的不变性，反过来由对应关系推知未知原象(即A，B，C两两联成的线性图形)也是一直线。

在上述推理中，映射就是射影变换，映象关系就是圆内接六边形及三双对边延长线的三交点。求得的映象就是三交点联成的一直线，反过来的对应关系便是反演。确定下来的未知原象就是由A，B，C联结成的直线图形。

例三 微积分学中寻求某些幂级数的和函数问题，也往往可根据RMI原则来获得解决。例如，试求下列幂级数的和函数：

$$S(x)=\frac{x^2}{1\cdot 2}-\frac{x^3}{2\cdot 3}+\frac{x^4}{3\cdot 4}-\cdots \quad (|x|<1).$$

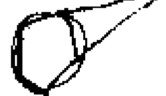
这里所谓求和函数，就是要求出和函数$S(x)$的初等表达式。因此，这种表达式就是我们所寻求的对象，即未知原象。当然，和函数实质上已被上述幂级数所确定，所以幂级数与$S(x)$的相等关系便是原象关系。我们知道，幂级数在它的收敛区域内可以逐项积分和逐项求微商。就目前问题而言，采用微商算子D所给出的如下映射

$$\varphi: \quad f\longmapsto Df\equiv\frac{d}{dx}f \ ,$$

可以把$S(x)$变换成一个较简单的熟知对象（映象）：

$$S'(x)=x-\frac{x^2}{2}+\frac{x^3}{3}-\cdots=\log(1+x).$$

在原象f满足$f(0)=0$的限制下，映射D是一对一的，其反演是



$$\varphi^{-1}: \quad g\longmapsto D^{-1}g\equiv\int_0^x g(t)dt.$$

因此，$S'(x)$的反演结果便给出了原象表示式

$$S(x)=\int_0^x \log(1+t)dt=(1+x)\log(1+x)-x.$$

上述解题过程显然符合RMI原则的框架。今再举一例，试求下列幂级数的和函数：

$$S(x)=1\cdot2+2\cdot3x+3\cdot4x^2+\cdots\quad(|x|<1)\ .$$

此时采用映射

$$\varphi:\ f(x)\longmapsto\int_0^x\int_0^v f(u)dudv,$$

可得

$$\int_0^x\int_0^v S(u)dudv=x^2+x^3+x^4+\cdots=\frac{x^2}{1-x}\qquad(|x|<1).$$

施行反演

$$\varphi^{-1}:\quad g(x)\longmapsto\frac{d^2}{dx^2}g(x),$$

则得

$$S(x)=\frac{d^2}{dx^2}\left(\frac{x^2}{1-x}\right)=\frac{2}{(1-x)^3}\qquad(|x|<1),$$

这便是所要寻找的原象的表示式.

上述诸例表明，应用RMI原则去处理数学问题时，重要关键是如何选取合适的映射．事实上，只有当映射选得好，才能使问题迎刃而解。

例四 熟知解析几何方法可用以处理平面几何问题，其基本思想方法也符合RMI原则。按照解析几何，笛卡儿坐标平面上的点用数偶（x，y）来表示，直线和圆分别用x，y的一次和二次方程式来表示。这样，作为原象的几何图形——点、线、圆等便和作为映象的（x，y）及含x，y的一次、二次式一一对应起来。这种对应关系即是映射关系。

通常，一个几何问题无非是关于某些特定几何图形间的关系问题。这种关系结构问题在上述对应（映射）下便转化为代数式之间的关系问题（即映射关系）．于是通过代数运算不难求得所需要的一些代数关系。这些关系再翻译回去就可得出原来几何图形（原象）间的某种几何结论，而这正是原来要求解答的问题。

上述思想方法可用框图表示如下：

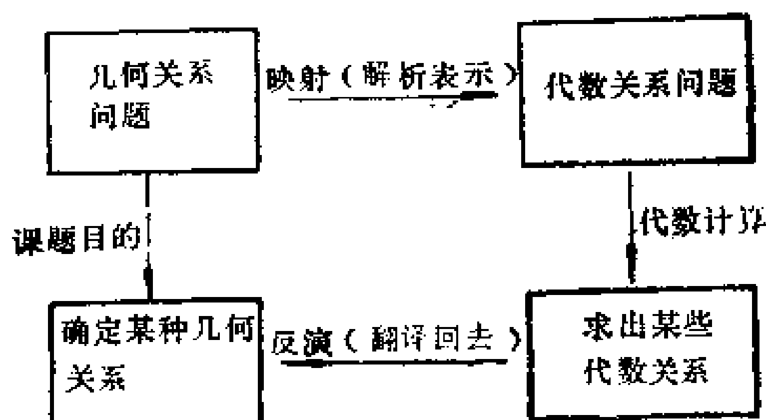


图3-5

今举一例，以示上述原则的具体运用。设H为三角形ABC的垂心，D，E，F为三高的垂足。自D分别作AB，BH，CH，CA的垂线，其垂足分别为K，L，M，N。求证此四点共线。

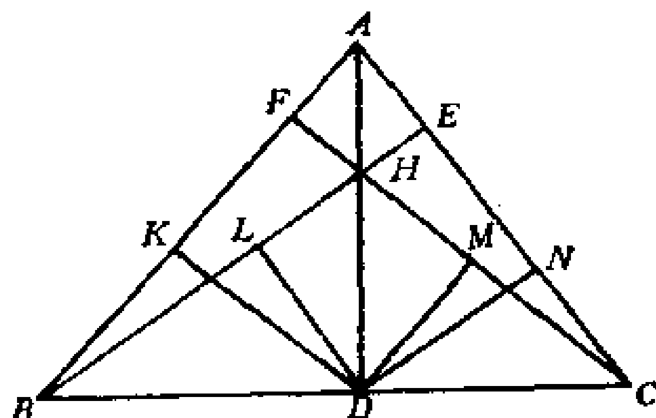


图3-6

为简化问题，取D为坐标原点，DC，DA分别为x，y轴。设置坐标A $(0, a)$，B $(b, 0)$，C $(c, 0)$。这样，便不难列出原象图形中十条直线在映象系统中所相应的十个一次方程式（为节省篇幅，方程具体形式不一一写出）。只需求解各对一次联立方程，便可求得各点坐标：

$$N\left(\frac{a^2c}{a^2+c^2}, \frac{ac^2}{a^2+c^2}\right), \quad K\left(\frac{a^2b}{a^2+b^2}, \frac{ab^2}{a^2+b^2}\right),$$

$$M\left(\frac{b^2c}{a^2+b^2},\ -\frac{abc}{a^2+b^2}\right),\quad L\left(\frac{c^2b}{a^2+c^2},\ -\frac{abc}{a^2+c^2}\right).$$

这些点的坐标分别满足解析几何中的三点共线条件

$$\begin{vmatrix} \dfrac{a^2c}{a^2+c^2} & \dfrac{ac^2}{a^2+c^2} & 1 \\ \dfrac{a^2b}{a^2+b^2} & \dfrac{ab^2}{a^2+b^2} & 1 \\ \dfrac{b^2c}{a^2+b^2} & -\dfrac{abc}{a^2+b^2} & 1 \end{vmatrix}=0$$

和

$$\begin{vmatrix} \dfrac{a^2c}{a^2+c^2} & \dfrac{ac^2}{a^2+c^2} & 1 \\ \dfrac{a^2b}{a^2+b^2} & \dfrac{ab^2}{a^2+b^2} & 1 \\ \dfrac{c^2b}{a^2+c^2} & -\dfrac{abc}{a^2+c^2} & 1 \end{vmatrix}=0 .$$

这两个代数关系式的几何解释是N，K，M共线和N，K，L共线。故结论是N，K，M，L四点共线。

例五 在上例中，如果把作为映象所在域的笛卡儿坐标平面换成高斯的复数平面，并对点、直线、圆等几何对象分别对应地表示成如下的复数形式：

几何点$\longleftrightarrow Z=x+iy$ $\quad(i=\sqrt{-1})$；

过Z_1，Z_2的直线$\longleftrightarrow Z-Z_1=\lambda(Z_2-Z_1)$ $\quad(-\infty<\lambda<\infty)$；

以Z_0为心以ρ为半径的圆$\longleftrightarrow |Z-Z_0|=\rho$。

在上述对应关系作成的映射下，凡平面几何中涉及点、线、圆的几何关系等问题，均可表成复数平面上的某些代数等式关系。这样，只需通过复数的一些代数计算，把得出的关系式重新翻译成几何关系，便可获得几何命题的证明。因此，如上所述的基本思想方法，仍可纳入RMI原则框架。在《初等数学论丛》第1辑有常庚哲同志写的《用复数解几何题》一文，文中不少有趣味

的例题，值得参考。

用代数计算或复数计算去代替几何命题的论证，方法比较机械，一般不会遇到险阻。这样做的好处是，可以避免在几何证题中为寻求特殊的“巧思”而耗费太多的时间。

§4 几个较难一点的例子

这里将介绍一种非常有用的映射，叫做**幂级数变换**。对于寻找或确定一个数列来说，通常又把它叫做“发生函数”。这是在近代组合数学及概率统计学中应用很广泛的一种映射方法。

方法的大意是：为了研究一个数列，例如$\{a_n\}$：$a_0, a_1, a_2, \cdots, a_n\cdots$的结构，我们利用一个形式幂级数

$$a_0+a_1t+a_2t^2+\cdots+a_nt^n+\cdots$$

与$\{a_n\}$对应起来。这样我们就把一个离散性的数列对象换成了一个在分析学上便于处理的解析对象。正因为在分析学中对于幂级数运算已有一套固定的方法，所以在上述对应（映射）下，作为幂级数系数的数列$\{a_n\}$的结构问题也就便于通过幂级数运算去分析研究了。

例一 试按照RMI原则求解差分方程

$$f_n=f_{n-1}+f_{n-2} \qquad (n=2, 3, 4, \cdots),$$

其初始条件为$f_0=1$，$f_1=2$。

差分方程就是数列间的“递推关系”。所谓解差分方程，就是要求根据递推关系及初始条件去确定整个数列$\{f_n\}$的构造形式，也即要求得出f_n的一般表达式。为此，可引入如下形式的映射（幂级数变换）：

$$\{f_n\}\longmapsto F(t)=1+\sum_{n=0}^{\infty}f_n\cdot t^{n+1}.$$

此处在幂级数中添加1作为常数项只是为了下面的计算方便，其

实，令f_0代表常数项，f_n表示t^n的系数，方法仍然可行。

在上述映射关系中，$\{f_n\}$是原象（未知元），$F(t)$是它的映象。因为$\{f_n\}$是未知的，所以$F(t)$也是未知的。但是根据给定的原象关系连同初始条件，可以相应地获得映象$F(t)$之间的关系：

$$
\begin{aligned}
F(t) &= 1+f_0t+f_1t^2+f_2t^3+\cdots+f_nt^{n+1}+\cdots \\
&= 1+t+f_0t^2+f_1t^3+\cdots+f_{n-1}t^{n+1}+\cdots \\
&\quad +t^2+f_0t^3+\cdots+f_{n-2}t^{n+1}+\cdots \\
&= 1+tF(t)+t^2F(t).
\end{aligned}
$$

这样，我们就从原象关系导出了映象关系

$$F(t)=1+(t+t^2)F(t).$$

由此求出映象

$$F(t)=1/(1-t-t^2).$$

从幂级数展开式求出系数序列的过程，叫做幂级数变换的反演。作为最后一步，还须对求得的映象$F(t)$作反演。事实上，利用分项分式法可得

$$F(t)=\frac{1}{1-t-t^2}=\frac{1}{a-b}\left(\frac{a}{1-at}-\frac{b}{1-bt}\right),$$

其中$1-t-t^2=(1-at)(1-bt)$，且

$$a=\frac{1}{2}\left(1+\sqrt{5}\right), \qquad b=\frac{1}{2}\left(1-\sqrt{5}\right).$$

因此可得出$F(t)$的展开式

$$F(t)=\sum_{n=0}^{\infty}\frac{a^{n+1}-b^{n+1}}{a-b}t^n=1+\sum_{n=0}^{\infty}f_n\cdot t^{n+1}.$$

比较即得

$$f^n=\frac{1}{\sqrt{5}}\left\{\left(\frac{1+\sqrt{5}}{2}\right)^{n+2}-\left(\frac{1-\sqrt{5}}{2}\right)^{n+2}\right\}.$$

这样便完全确定了$\{f_n\}$的具体结构形式。

值得提及的是，f_n（或记作F_{n+1}，$n=0, 1, 2, \cdots$）实际就是优选法理论研究中经常用到的斐波那奇（Fibonacci）数列，

俗名"兔子数列"。

例二 一般的二项系数记为

$$\binom{\alpha}{k}=\frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!},$$

这里的α可以是任意实数。特别$\binom{0}{0}=1$。今设α和β是任意两个实数，试求下式之和

$$c_n=\sum_{k=0}^{n}\binom{\alpha}{k}\binom{\beta}{n-k}.$$

这里所谓"求和"的意思是要求找出c_n的简明表达式，即不再包含求和运算的表达式。为此目的，我们采取如下步骤：

第一步，明确原象关系。简记

$$a_n=\binom{\alpha}{n}, \qquad b_n=\binom{\beta}{n},$$

则所求对象$\{c_n\}$就由如下关系所确定：

$$c_n=\sum_{k=0}^{n}a_k\cdot b_{n-k} \qquad (n=0,1,\cdots),$$

这就是原象关系，其中尚未找出一般简明表示式的c_n就是未知元。

第二步，引用幂级数变换作为映射工具：

$$\{a_n\}\longmapsto A(t)=\sum_{n=0}^{\infty}a_nt^n;$$

$$\{b_n\}\longmapsto B(t)=\sum_{n=0}^{\infty}b_nt^n;$$

$$\{c_n\}\longmapsto C(t)=\sum_{n=0}^{\infty}c_nt^n.$$

第三步，找出映象关系。由幂级数的乘法规则和原象关系，可得

$$C(t)=\sum_{n=0}^{\infty}c_n t^n=\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}a_k\cdot b_{n-k}\right)t^n=A(t)\cdot B(t).$$

这说明映象关系就是简单的乘积关系。

第四步，确定$\{c_n\}$的映象。根据牛顿二项式定理，

$$A(t)=\sum_{n=0}^{\infty}\binom{\alpha}{n}t^n=(1+t)^{\alpha},$$

$$B(t)=\sum_{n=0}^{\infty}\binom{\beta}{n}t^n=(1+t)^{\beta}\qquad(|t|<1).$$

从而立即得出$\{c_n\}$的映象为

$$C(t)=(1+t)^{\alpha}\cdot(1+t)^{\beta}=(1+t)^{\alpha+\beta}.$$

最后一步，作反演。求$C(t)$的展开式系数，可得

$$c_n=\binom{\alpha+\beta}{n}.$$

这样便把原象确定下来。结论是下列组合恒等成立：

$$\sum_{k=0}^{n}\binom{\alpha}{k}\binom{\beta}{n-k}\equiv\binom{\alpha+\beta}{n}.$$

这就是组合数学中著名的范德蒙（Vandermonde）定理。

从例一我们已经看到幂级数变换在解差分方程时的作用。十九世纪初期的数学家，例如拉普拉斯（Laplace），就已经领悟到差分方程不过是微分方程的离散化形式。由此通过对比联想，拉普拉斯成功地引入积分变换

$$f(t)\longmapsto F(s)=\int_0^{\infty}f(t)e^{-ts}dt$$

作为求解微分方程的映射工具。这一映射工具非常有效，在现代数学文献及教本中称之为拉氏变换，它的研究已成为现代运算微积分的理论基础。

为说明上述映射的功能，这里举一个《工程数学》教本中的简单例子。

例三 试按照RMI原则求解微分方程

$$y'' + 2y' - 3y = e^{-t},$$

其初始条件为$y(0)=0$，$y'(0)=1$。

第一步，明确原象关系。在本题中未知原象为$y(t)$，它所满足的微分方程和初始条件组成原象关系。

第二步，应用拉氏变换作映射。原象$y(t)$的映象是

$$Y(s) = \int_0^{\infty} e^{-st} y(t)dt \text{ 。}$$

今对微分方程两边的函数同时作拉氏变换，并顾及初始条件。利用初等微积分中的分部积分法，不难求得映象关系

$$s^2Y(s) - 1 + 2sY(s) - 3Y(s) = \frac{1}{s+1}.$$

第三步，从映象关系求出映象。可得

$$\begin{aligned} Y(s) &= \frac{s+2}{(s+1)(s-1)(s+3)} \\ &= \frac{3}{8(s-1)} - \frac{1}{4(s+1)} - \frac{1}{8(s+3)}. \end{aligned}$$

第四步，求$Y(s)$的逆变换（即拉氏变换的反演）。利用拉氏变换对照表，可以立即写出$Y(s)$的原象：

$$y(t) = \frac{1}{8}(3e^{t} - 2e^{-t} - e^{-3t}).$$

这便是微分方程的解。

例一、例三中的求解方法可以推广到一般情形，即幂级数变换与拉氏变换可分别用作求解高阶常系数差分方程与微分方程的映射工具。欲知其详，请读者参考差分方程与微分方程的有关书籍。

近代概率论中常用的特征函数方法以及关于一些数学物理方程（偏微边值问题）通过变分原理化为积分方程极值问题而进行求解的方法等，显然都可归入RMI原则的具体应用范围中去。还有关于积分方程的Fredholm基本公式也可根据RMI原则来导出，

但须利用从连续结构到离散结构的映射方法。

§5　用RMI原则分析“不可能性命题”

数学上有一类问题，叫做**不可能性问题**，即关于某类数学对象不可能存在某种属性，或者某个系统不可能存在某种元素的那些问题。所谓不可能性命题就是否定某种存在性的命题。例如，$\sqrt{2}$不是有理数（即有理数集合中不包含$\sqrt{2}$）以及“三等分角的尺规作图不可能性”等等都是不可能性命题。

RMI原则不仅可用以直接探索某些关系结构的未知原象，而且也可用以分析论证某些不可能性命题。

例一　试按RMI原则证明实数集合的不可数性（即用自然数编号不可能把全体实数都排列成一个无穷数列，如x_1，x_2，…，x_n，…）。

我们只需证明（0，1）内的全体实数点不能排成一个数列（点列）$\{x_n\}$：x_1，x_2，x_3，…，x_n，…。为此，又只需证明（0，1）内的任何一个处处稠密的无穷数列$\{x_n\}$都不能完全取遍（0，1）内的全体实数（即点列$\{x_n\}$不能填满（0，1）区间）。

今设$\{x_n\}$是一个各数皆相异而处处稠密于（0，1）的任何数列。我们要证明$\{x_n\}$决不可能充满（0，1）区间。也即（0，1）内总有某些实数不包含在数列$\{x_n\}$之内。为此，我们又只须证明$\{x_n\}$内总存在某个子序列$\{x_{n_k}\}$：x_{n_1}，x_{n_2}，…，x_{n_k}，…，它的极限值含于（0，1）内但却不在数列$\{x_n\}$之内。可是这样的子序列是否存在还不知道。下面我们就根据RMI原则来推断这种子序列的存在性。

我们把具有上述性质的子序列称为$\{x_n\}$的特异子序列。这种子序列乃是包含在$\{x_n\}$序列结构中的未知原象（连存在性也不知道的原象！）。

第一步，对$\{x_n\}$作一个保序有理映射：

$$x_n \longmapsto y_n, \quad 0<y_n<1.$$

这个映射的具体作法是：另取一个数轴的单位区间（0，1）作为映象点列$\{y_n\}$的所在区间。首先取（0，1）的中点作为y_1（即$y_1=\frac{1}{2}$）。如果$x_2<x_1$则命y_2为（0，y_1）的中点，否则命y_2为（y_1，1）的中点。这样，y_1，y_2间的位置排序关系和x_1，x_2间的排序一致。一般说来，映射的构造规则是：（1）要求映射是保序的，即对应于$x_i<x_j$应有$y_i<y_j$；（2）如果按序号来讲，x_k是继x_i，x_j二点之后（即当$k>i$，$k>j$时）最先落入空白区间（x_i，x_j）的点，意即假设（x_i，x_j）中不含有编号小于i，j的x点，而$x_i<x_k<x_j$，且k是满足条件$k>i$，$k>j$的最小序号，则规定$y_k=\frac{1}{2}(y_i+y_j)$，从而仍有保序关系$y_i<y_k<y_j$。

按上述映射规则，可见每个y_n均可表成形如$m/2^n$的有理分数（也即二进小数）。

第二步，确定映象数列的具体结构。为此，我们应证明下述引理。

引理 数列$\{y_n\}$取遍形如$m/2^k$的全体有理数，其中$m=1$，2，…，2^k-1；$k=1$，2，3，…。特别，$y_1=1/2$相当于$k=1$，$m=1$。

可用反证法证明这个引理。假定$\{y_n\}$并非取遍一切形如$n/2^k$的有理数，则必存在某个最小的$k=k_0$（$\geqslant 2$）和某个奇数$m=2j+1$致使$m/2^{k_0}$不包含在数列$\{y_n\}$内。既然k_0是最小者，故这样的$m/2^{k_0}$必处于$\{y_n\}$中某两数（两点）

$$y_\nu=j/2^{k_0-1} \text{ 与 } y_\mu=(j+1)/2^{k_0-1}$$

$$(0\leqslant j<2^{k_0-1}-1)$$

之中值位置。但由$\{x_n\}$的稠密性，必有某个x_λ首先落入区间(x_ν，x_μ）中。于是对应于x_λ应有

$$y_\lambda=\frac{1}{2}(y_\nu+y_\mu)=m/2^{k_0},$$

这是含于$\{y_n\}$内的一个数。由此得一矛盾，引理得证。

第三步，确定$\{y_n\}$中的一个特异子序列。含于(0，1)内而不在$\{y_n\}$中的实数是很多的。例如$y^*=1/3$就不在$\{y_n\}$内。显然，$\{y_n\}$内存在一个单调增序列y_{n_k}和一个单调降序列$\{y_{m_k}\}$，使得

$$\lim_{k\to\infty} y_{n_k}=y^*=\lim_{k\to\infty} y_{m_k},$$

此处， $y_{n_k}<y_{m_k}$， $k=1, 2, \cdots$。

第四步，通过反演获得$\{x_n\}$的特异子序列。事实上，由逆映射我们得到

$$x_{n_k}<x_{m_k}, \quad k=1,2,\cdots.$$

$$\lim_{k\to\infty} x_{n_k}=x_1^*\leqslant x_2^*=\lim_{k\to\infty} x_{m_k}.$$

易知$x_1^*=x_2=x^*$。否则，将有x_r含于(x_1^*, x_2^*)内，从而由保序关系有y_r含于一切(y_{n_k}, y_{m_k})内，这样，$y_r=y^*$，导致矛盾。进一步，可推知x^*不在$\{x_n\}$内，否则将有

$$x^*=x_r\longleftrightarrow y_r=y^*=\frac{1}{3},$$

又导致矛盾。因此，$\{x_{n_k}\}$和$\{x_{m_k}\}$都是特异子序列。

既然x^*含于（0，1）而不在$\{x_n\}$内，故已证明$\{x_n\}$不能填满区间（0，1）。实数集合的不可数性至此获证。

学过分析数学的人都知道，应用康托（Cantor）的“对角线论证法”可以很简短地证明实数集合的不可数性。但上面的证法符合RMI原则的框架，看来还是比较自然的，直观的；因为它通过映象数列$\{y_n\}$不能填满区间（0，1）的事实，直接对应地导出了原象数列$\{x_n\}$不能填满（0，1）的结论。

例二 试应用RMI原则概略地分析几何三大难题尺规作图解法的不可能性。

所谓几何三大难题，就是古希腊时代提出的三个尺规作图

题：

(1) 要求作一正方形，使其面积等于单位圆的面积π，也即要求作出长为$\sqrt{\pi}$的线段（这叫做“圆化方问题”）。

(2) 要求作一立方体，使其体积等于原单位立方体体积的二倍，也即要求作出长度为$\sqrt[3]{2}$的线段来（这叫做“倍立方问题”）。

(3) 要求把任意角进行三等分（这叫做“三分角问题”）。

利用任何有限个步骤的尺规作图法都不可能解决上述三题一事，早在上一世纪就已获得彻底澄清。但是关于“不可能性”的论证细节很繁长（需要用到伽罗瓦理论等知识），这里只能对其基本思路作一个十分简要的描述性说明。

在如上(1)(2)两题的陈述中，把作图要求转化为求作长度为$\sqrt{\pi}$与$\sqrt[3]{2}$的线段，这实际上已经用到了形数对应法则。事实上，为了确定尺规作图可能性准则，必须应用解析几何的形数对应（映射）方法才能获得成功。

分析一切几何作图过程的基本步骤，无非是要确定一些点的位置。在笛卡儿坐标平面上，点的位置由纵横坐标所确定。因此，作图过程的基本步骤又可归结为求作定长的线段。定长线段乃是最基本的几何量。

任何作图不能从“乌有”开始，因此，不失一般性，我们总可把作图开始用的原来给定的几何量看作一个单位长线段。单位长是人为规定的，作此规定对以后的分析叙述比较方便。下面我们采用解析几何学的形数对应观点分析问题。

如大家所知，一旦取定单位长线段对应于数量1之后，数轴上那些代表有理数的坐标点，便可由尺规作出。如果考虑笛卡尔坐标平面，则一切以有理数为纵横坐标的位置点便都可由尺规作出。这些点可称为有理点。用直尺作图可以联结任意两个有理点，当然还可以借助于圆规作出各个有理数的平方根以及平方根的平方根等数量。

在上述坐标平面上，使用直尺作出的直线，使用圆规作出的圆周，它们的方程式无非是如下的形式：

$$ax+by+c=0;$$

$$(x-d)^2+(y-e)^2=r^2,$$

这里a，b，c，d，e，r等也都是些有理数经过有限次+、-、×、÷、$\sqrt{\ }$等五则运算得出的数量。所谓尺规作图法，按形数对应的解析几何观点来看，无非是利用直线与直线相交、直线与圆周相交、圆与圆相交等截取交点的几种基本方式来进行的。自然，这些交点坐标相应地是由两个一次联立方程、一次与二次联立方程、两个二次联立方程的代数解法来确定的。因此，所能得出的数量都是由有理数经过有限多次五则运算表示出的数量。尺规作图所能作出的仅限于这类形式的数量（通常表现为线段长的几何量）。这便是尺规作图法的一条准则，它可以用来判断哪些作图题可用尺规解决，哪些则不能解决。

根据形数对应法则，我们已经知道倍立方问题与圆化方问题所要求作的数量分别是$\sqrt[3]{2}$和$\sqrt{\pi}$。可是，根据伽罗瓦理论和1882年林德曼（Lindemann）已证明π与$\sqrt{\pi}$都是超越数的事实，我们知道$\sqrt[3]{2}$与$\sqrt{\pi}$都不属于上述尺规作图准则所规定的数量范围。因此，问题（1）和（2）从根本上说来，都是尺规作图法不可能解决的问题。

为说明三分角问题不能解，只需以60°角为例说明即可。为把60°三等分，必须用尺规作出数量cos20°或sin20°。于熟知的三角恒等式

$$\cos 3x=4\cos^3 x-3\cos x$$

中置$x=20°$，则$\cos 3x=\cos 60°=\frac{1}{2}$，故上述等式变为$y=\cos 20°$的三次方程

$$4y^3-3y-\frac{1}{2}=0.$$

利用三次方程的求根知识，可以证明这个方程中的y有一个正实根，两个负实根，它们都必须用有理数的立方根表示出来，而无法表为尺规作图准则中的数量形式。因此，即使是60°的三分角问题也不能解。

从给定的初始线段出发，凡能由尺规作图得出的一切线段，统称为可作几何量。这些量构成一个类。几何三大难题要求作出的线段可称为待作几何量。在形数对应法则下它们所对应的解析量分别为$\sqrt{\pi}$，$\sqrt[3]{2}$和cos20°等等。于是以上所述尺规作图不可能性的推理过程显然可表示为下列框图。

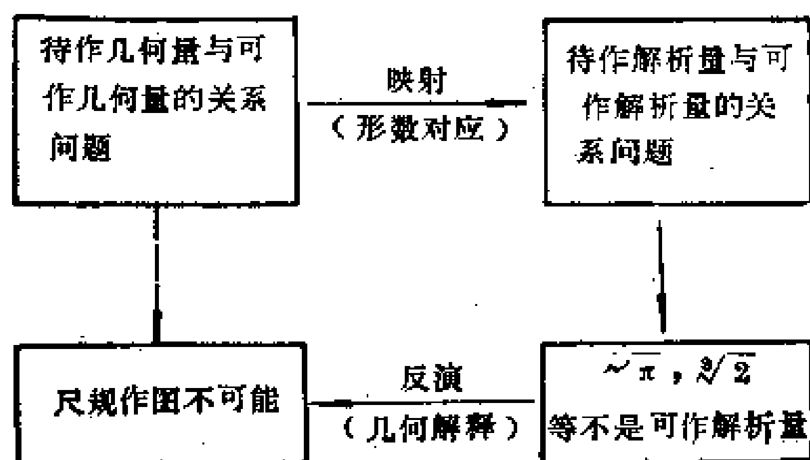


图3－7

这说明关于作图问题不可能性的分析也可纳入一般RMI原则的范畴。关于几何作图理论的详细讨论，请参考李克正同志《几何作图杂谈》一文，见《初等数学论丛》第1辑。

§6　关于RMI原则的补充说明

我们已经举了十个例子，阐明RMI原则的种种应用。但要注意，“RMI”并不是一个万能的原则。事实上有许多数学问题是不能依靠“RMI”来解决的。RMI原则的主要作用，只是帮助启发解决问题的思路而已。

在我们列举的一些例子中已经看到，对于特定的数学问题，为了有效地参照RMI原则所规划的框架去解决问题，必须利用可定映映射去完成定映这一步骤。

"定映"，就是要从映象关系里确定出未定目标映象的属性（数量、形式或某些性质）。这在不同类型的数学问题里常常表现出不同程度的难度。有时甚至需要引用若干重要定理才能完成这一步。例如，在§5（例二）的分析中我们就看到这种情形。

"反演"，就是逆变换或逆映射。如果仅从对应观点看，这并不会带来任何困难，但如果处理的是解析变换的逆变换（例如拉氏变换的逆变换等等），则需要寻找合适的反演公式等等。这些也是数学领域里的研究课题。

凡是利用可定映映射φ联系起来的原象关系结构S与映象关系结构S^*便称为**可定映系统**。特别，如果逆映射φ^{-1}具有某种能行性，即能将目标原象的某种所需要的性状经有限步确定下来者，则称该系统为**可解结构系统**，可简记为$(S, S^*; \varphi)$。

作成可解结构系统的基本关键在于可定映映射的存在性。根据对前面讲过的一系列例子的分析和概括，还可对可定映这一概念作更为确切的说明。为此，先解释下述名词：

数学手续——凡是由数值计算、代数计算、解析计算（包括极限手续等）、逻辑演算以及数学论证等步骤作成的形式过程都称之为数学手续。

于是，对于给定的一个具有目标原象x的关系结构S，如果有这样的一个可逆映射φ，它将S映成映象关系结构S^*，在S^*中通过某种形式的有限多步数学手续，能把目标映象$x^*=\varphi(x)$的某种所需要的性状确定下来的话，那么就称φ为可定映映射。特别，如果φ^{-1}还具有能行性，则获得一个可解结构系统$(S, S^*; \varphi)$。

最后，我们可将RMI原则重新表述如下：对于含有某种目标原象的原象关系结构S，先设法寻找一个可定映映射φ，同时考虑到φ的逆映射φ^{-1}具有合乎问题需要的能行性。于是通过"关系—

映射—定映—反演”诸步骤便可把所要求的目标原象确定下来。这个一般性方法原则就称为数学上的RMI原则。

数学上的RMI原则对数学工作者很是有用。小而言之，可利用该原则解决个别的数学问题。大而言之，甚至可以利用该方法原则作出数学上的重要贡献。一般说来，如果谁能对一些十分重要的关系结构S，巧妙地引进非常有用且具有能行性反演φ^{-1}的可定映映射φ，就能作出较重要的贡献。数学史上的纳皮尔(Napier)，拉普拉斯（Laplace），弗雷德霍姆（Fredholm）等就是一些突出的例子。

第4讲 略论数学公理化方法

§1 公理化方法的意义和作用

公理化方法在近代数学的发展中起过巨大的作用，可以说，它对各门现代数学都有极其深刻的影响。即使在数学教学中，公理化方法也是一个十分重要的方法。

所谓**公理化方法**（或公理方法），就是从尽可能少的无定义的原始概念（基本概念）和一组不证自明的命题（基本公理）出发，利用纯逻辑推理法则，把一门数学建立成为演绎系统的一种方法。所谓基本概念和公理，当然必须反映数学实体对象的最单纯的本质和客观关系，而并非人们自由意志的随意创造。

如所共知，希尔伯特1899年出版的《几何学基础》一书是近代数学公理化的典范著作。该书问世后的二、三十年间曾引起西方数学界的一阵公理热，足见其影响之大。希尔伯特的几何公理系统实际是在前人的一系列工作成果基础上总结出来的，书中的公理条目也曾屡经修改。直到1930年出第七版时，还作了最后修改。这说明一门学科的公理化未必是一次完成的，公理化过程可以是包含着一些发展阶段的。

谈到数学公理化的作用，至少可以举出如下三点：(1)这种方法具有分析、总结数学知识的作用。凡取得了公理化结构形式的数学，由于定理与命题均已按照逻辑演绎关系串联起来，故使用起来也较方便。(2)公理化方法把一门数学的基础分析得清清楚楚，这就有利于比较各门数学的实质性异同，并能促使和推动新理论的创立。(3)数学公理化方法在科学方法论上有示范作用。这种方法对现代理论力学及各门自然科学理论的表述方法

都起到了积极的借鉴作用。例如，本世纪四十年代波兰的巴拿赫（Banach）曾完成了理论力学的公理化；物理学家还把相对论表述为公理化形式，等等。

§2　公理化方法发展简史

分析公理方法的历史发展，大致可分成三个阶段：

一是公理方法的产生阶段　大约在公元前三世纪，希腊的哲学家和逻辑学家亚里斯多德（Aristotle）总结了古代积累起来的逻辑知识，以演绎证明的科学（主要是数学）为实例，把完全三段论作为公理，由此推导出别的所有三段论法（共分了十九个格式）。因此可以认为，亚里斯多德在历史上提出了第一个成文的公理系统。

亚里斯多德的思想方法深深地影响了公元前三世纪的希腊数学家欧几里得，后者把形式逻辑的公理演绎方法应用于几何学，从而完成了数学史上的重要著作《几何原本》。欧几里得从古代的量地术和关于几何形体的原始直观中，用抽象分析方法提炼出一系列基本概念和公理。他总结概括出14个基本命题，其中有5个公设和9条公理。由此出发，他运用演绎方法将当时所知的几何学知识全部推导出来。这便是古代数学公理化方法的一个辉煌成就。

《几何原本》的问世标志了数学领域中公理方法的诞生。它的贡献不在于发现了几条新定理，而主要在于它把几何学知识按公理系统的方式妥切安排，使得反映各项几何事实的公理和定理都能用论证串联起来，组成了一个井井有条的有机整体。

二是公理方法的完善阶段　如所知，欧氏几何的公理系统是不够完善的。其主要的不足之处可以概括为：（1）有些定义是不自足的，亦即往往使用一些未加定义的概念去对别的概念下定义。（2）有些定义是多余的，略去它毫不影响往后的演绎和展

开。（3）有些定理的证明过程往往依赖于图形的直观。

事实上，几何原本的不足之处早已为古代学者所觉察，例如，Archimedes为严格表述有关长度、面积和体积的测量理论，就曾对欧氏几何公理系统作过必要的扩充，人所共知的Archimedes公理（任给非负数量a和b，$a<b$，存在n使得$na>b$）便是其中一例。可以说，自从Archimedes以后，人们一直在努力完善几何原本的陈述。

另一方面，由于第五公设（即平行线公理）在陈述与内容上的复杂和累赘，古代学者们早就怀疑地指出，第五公设是不是多余的，它能否从其他公设、公理中逻辑地推导出来？而且进一步认为，欧几里得之所以把它当作公设，只是因为他未能给出这一命题的证明。因而，学者们纷纷致力于证明第五公设。但是所有试证第五公设的努力均归于失败，在这些失败之中唯一引出的正面结果便是一串与第五公设相等价的命题被发现。

据说在欧几里得以后的两千多年时间里，几乎难以发现一个没有试证过第五公设的大数学家，其中特别著名的可以提到Legendre、Saccheri和Lambert。不妨把他们处理第五公设问题的思想方法作一简略的介绍。

如图4－1所示，Lambert从讨论四角形$ABCD$出发，在这个四角形中，有三个内角被假定为直角d，我们把余下的那个没有假定为直角的内角叫做Lambert角并记为$l_{d}^{<}$。

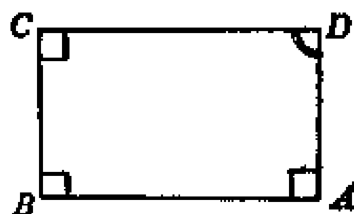


图4－1

如图4－2所示，Saccheri从讨论四角形$AA'B'B$出发，在这四角形中，假定$AA'=BB'$，而夹着AB边的两个内角都是直角d，由图形的对称性易知该四角形的另外两个内角是相等的，人们称之为Saccheri角并记为$S_{d}^{<}$。

Legendre则从讨论三角形内角和$\Sigma(\triangle)$入手，对于$\Sigma(\triangle)$而言，显然有且仅有$\Sigma(\triangle)>2d$，$\Sigma(\triangle)=2d$和$\Sigma(\triangle)<2d$

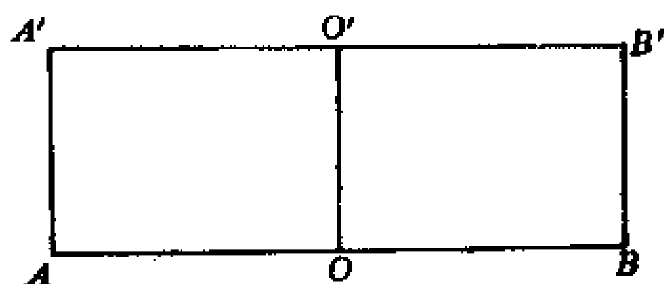


图4-2

三种可能。同样对于$S_a^{\leq}$来说，也有且仅有$S_a^{\leq}<d$、$S_a^{\leq}=d$和$S_a^{\leq}>d$三种可能；而对于$L_a^{\leq}$也是一样，或者$L_a^{\leq}<d$、或者$L_a^{\leq}=d$、或者$L_a^{\leq}>d$。他们分别证得了当设$\Sigma(\triangle)>2d$、$S_a^{\leq}>d$和$L_a^{\leq}>d$时必然导致矛盾，又分别证得了$\Sigma(\triangle)=2d$、$S_a^{\leq}=d$和$L_a^{\leq}=d$均与第五公设等价。于是，一旦能不以第五公设为基础而否定$\Sigma(\triangle)<2d$、$S_a^{\leq}<d$、$L_a^{\leq}<d$之一为真，第五公设便获证。他们都想用反证法来实现这一目的，亦即各想在$\Sigma(\triangle)<2d$、$S_a^{\leq}<d$或$L_a^{\leq}<d$的假定下导致矛盾。但是，他们除了成批地获得那些不合常情的定理（实质上就是属于非欧几何的定理）外，始终未能引出矛盾来。

基于两千多年来在证明第五公设的征途上屡遭失败的教训。十九世纪俄国年轻数学家Лобачевский产生了与前人完全不同的信念：首先，认为第五公设不能以其余的几何公理作为定理来证明；其次，除掉第五公设成立的欧几里得几何之外，还可以有第五公设不成立的新几何系统存在。于是，他在剔除第五公设而保留欧氏几何其余公理的前提下引进了一个相反于第五公设的公理："过平面上一已知直线外的一点至少可以引两条直线与该已知直线平行"。这样，Лобачевский在与前人完全不同的思想方法基础上构造了一个新的几何系统，它与欧几里得几何系统相并列。后来，人们又证明了这两个部分地互相矛盾的几何系统竟是相对相容的（参阅本讲§5），亦即假定其中之一无矛盾，则另一个必定无矛盾。如此，只要这两个系统是无矛盾的，第五公

设与欧氏系统的其余公理就必定独立无关。现在人们就用Лобачевский的名字对这一新几何命名，并把一切不同于欧氏几何公理系统的几何系统统称为非欧几何。应当指出，独立地发现这个新几何的还有大数学家高斯和匈牙利青年大学生Bolyai。但是高斯由于害怕学术界顽固守旧势力的攻击而始终不敢公开发表他的结果。

非欧几何学中的一系列命题都和人们的朴素直观不相符合。这是它在开创阶段之所以遭受人们冷嘲热讽的重要原因。但是，这种背离直观的几何学在逻辑系统内没有矛盾，演绎论证的严格性也是无暇可击的。事实上，非欧几何给人们开拓了"空间"的概念（如所知，非欧几何的重要分支"黎曼几何"在爱因斯坦1915年创立"广义相对论"后，已得到了证实和应用）。非欧几何的产生，不仅为公理化方法进一步奠定了基础，而且为公理方法可以推广和建立新的数学理论提供了依据。

非欧几何的创立大大地提高了公理方法的信誉。接着便有许多数学家致力于公理方法的研究。例如，1871～1872年间德国数学家康托（Cantor）与戴德金（Dedekind）不约而同地拟成了连续性公理。德国数学家巴许（Pasch）在1882年拟成了顺序公理。正是在这样的基础上，希尔伯特于1899年发表了《几何学基础》一书，终于解决了欧氏几何的欠缺问题，完善了几何学的公理化方法。此书也就成为近代公理化思想的代表作。

三是公理方法的形式化阶段 欧氏《几何原本》表现的公理化可称之为"实体公理化"，因为在这样的公理系统中，概念直接反映着数学实体的性质，而且那些概念、定义、公理和论证的表述往往束缚于直觉观念的指导。但在希尔伯特于其《几何学基础》一书中对欧氏系统加以完善化以后，不仅在公理的表述或定理的论证中摆脱了空间观念的直觉成份，而且给出和奠定了对一系列几何对象及其关系进行更高一级抽象的可能性和基础。就是说，人们可以在高度抽象的意义下给出公理系统，只要能满足系

统中诸公理的要求，就可以使得该公理系统所设计的对象是任何什么事物，并且在公理中表述事物或对象之间的关系时，也可以具有其具体意义的任意性。这样，自从《几何学基础》问世之后，不仅公理化方法进入数学的其它各个分支，而且把公理化方法本身推向了形式化的阶段。

后来，公理化方法形式化之所以能取得成功，在很大程度上得力于康托创始的抽象集合论。如果没有集合论思想方法和数理逻辑学的近代发展，形式公理化方法也不可能获得新的进展。如所知，希尔伯特后来从事“元数学”即“证明论”的研究，又把公理方法推向一个新的阶段，即纯形式化阶段（请参阅第10讲有关内容）。其基本思想就是采用符号语言把一个数学理论的全部命题变成公式的集合，然后证明这个公式的集合是无矛盾的。详言之，在这个集合中凡定义、公理及定理均写成公式的形状，而定理的证明步骤也无非是一串符号公式作成的系列，系列中的最后一式即所要证明的结论。形式化公理方法不仅推动着数学基础研究，而且还推动着现代算法论研究，从而为数学应用于电子计算机等现代科学技术开辟了新的前景。

§3　公理化方法的基本内容

如前所述，数学公理化的目的就是要把一门数学表述为一个演绎系统。这个系统的出发点就是一组基本概念和公理。因此，如何引进基本概念和确立一组公理便是运用公理化方法的关键，也即这种方法的基本内容。

基本概念既是不加定义的概念，它们就必须是真正基本的，而无法用更原始更简单的概念去界定的概念。换言之，基本概念应该是最原始最简单的思想规定，它们必须是对数学实体的高度纯化的抽象。当基本概念确定以后，重要的问题是如何设置公理的问题。

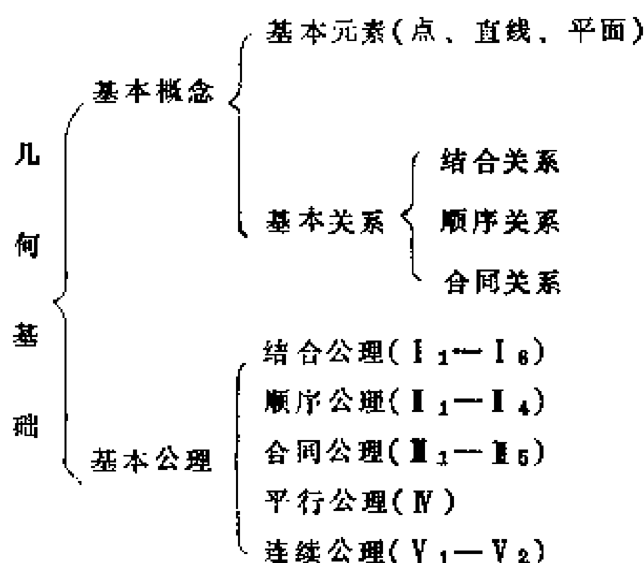
公理是对诸基本概念（例如基本元素、基本关系等概念）相互关系的规定。这些规定必须是必要的、合理的。详细说来，公理的选取和设置必须符合三条要求：一是**协调性要求**。协调性又称无矛盾性或相容性。这一要求是指在公理系统内，不允许同时能证明某一定理及其否定理。反之，如果能从该公理系统导出命题A和否命题“非A”（记作$\neg A$），则A与$\neg A$的并存便称之为矛盾。因此，无矛盾性要求是对公理系统的一个基本要求。二是**独立性**要求。这就是要求公理的数目减少到最低限度，不容许公理集合中出现多余的公理。因为多余的公理总可作为定理推证出来，又何必再把它列为公理呢？三是关于公理系统的**完备性**要求。这就是要确保从公理系统能够导出所论数学某分支的全部命题。因此，必要的公理不能省略，否则将得不到由它所能推得的结果。

一般说来，当一个公理系统满足上述三条要求时，即可认为是令人满意的系统了。但针对一个较复杂的公理系统要逐一验证三条要求，却并不是轻而易举的事，甚至至今不能彻底实现。例如，后面我们所要讨论的几何学公理系统，至今还只能在相对相容的意义下来论证它的无矛盾性等等。

§4　重要例子——几何学公理化方法

前面已经说过，人类从事几何学公理化工作，经历的年代最久，费去的精力最多，而获得的成果也最具有典型性，其影响自然也最深远。因此，这里专门来介绍一下几何学公理化的主要内容是很有意义的。

希尔伯特在所著《几何学基础》中引进的基本概念包括基本元素和基本关系，引进的基本公理共分五组二十条，即如下表所示。



这里的关键在于基本概念和公理的选取。对它们虽然不加定义和不加证明，但其选取并不是人为的、任意的。因为几何学的对象实体毕竟来源于现实世界，所以几何学的基本概念和公理必须符合客观实际。否则任凭主观随意创造，搞出来的“几何”就没有意义了（当然，非欧几何的意义不同，又当别论）。

希尔伯特《几何学基础》中的五组公理都是以几何实体为背景抽象概括出来的，所以显然是符合客观实际的。今将该书的五组公理逐一介绍如下（注意其中点、线、面、角等等都是无定义的抽象代名词）：

Ⅰ．**结合公理**

I_1：过A和B两点有一直线a。

I_2：过A和B两点至多有一直线a。

I_3：直线上至少有二点。又至少有三点不在同一直线上。

I_4：过不在同一直线上的三点A，B，C必有一平面α。在每一平面上至少有三点。

I_5：过不在同一直线上的三点A，B，C至多有一个平面。

I_6：如果一直线的两点在一平面α上，则该直线的每一点都在α上。

I_7：如果有两平面α，β有一公共点A，则它们至少还有另一

个公共点B。

I_8：至少有四点不在同一平面上。

Ⅱ. 顺序公理

II_1：若一点B位于点A与点C之间，则A，B和C为一直线上的三个点，且B也位于C与A之间。

II_2：至少有一点B位于任二点A与C作成的直线AC上，且C位于A，B之间。

II_3：直线上的任意三点中，至多有一点位于其它两点之间。

（注意。公理II_2和II_3隐含了直线的无限性。）

II_4（Pasch公理）：令A，B和C三点不在同一直线上，又设直线a位于A，B，C三点的平面上但不通过A，B或C。如果a穿过AB截段中的一个点，则a必穿过截段AC或BC中的一个点。

（注意。这条公理中用到"截段"这一概念，截段的定义是：设A，B为直线a上的两点，则点偶AB或BA就称为截段。所有A，B之间的点就称为截段内的点。A与B叫作截段的端点，而直线a上所有其它诸点称为截段外的点。）

Ⅲ. 合同公理

III_1：如果A，B为一直线a上的两个点，A'为直线a上或另一直线a'上的一个点，则在A'的给定一侧必可在a或a'上找到一点B'使得截段$A'B'$合同于AB。可用符号记为$AB\equiv A'B'$。

III_2：若$A'B'$和$A''B''$都与AB合同，则$A'B'\equiv A''B''$。

III_3：令AB与BC为直线a上无公共内点的两个截段，又令$A'B'$与$B'C'$为直线a'上无公共内点的截段。如果$AB\equiv A'B'$，且$BC\equiv B'C'$，则必有$AC\equiv A'C'$。

（注意。III_2与III_3把欧氏几何中原来泛指的等量公理进一步精确化了。）

III_4：令$\angle(h,k)$为平面α上二直线（射线）构成一个角，又令a'是位于平面α'上的一直线且给定a'在α'上的一侧，设h'为

a'的由点O'出发的一射线，则在a'上恰有一射线k'使得$\angle(h,k)$合同于$\angle(h', k')$且角$\angle(h', k')$的所有内点均位于a'的给定一侧。各个角均与其自己合同。

Ⅲ$_5$：关于两个三角形ABC和$A'B'C'$，如果有$AB \equiv A'B'$，$AC \equiv A'C'$且$\angle BAC \equiv \angle B'A'C'$，则必有$\angle ABC \equiv \angle A'B'C'$。

（注意．Ⅲ$_5$可用以推证$\angle ACB \equiv \angle A'C'B'$。）

Ⅳ．**平行公理**

令a为一直线而点A不在a上。则在a与A的平面上至多有一直线通过A而与a不相交。

（注意．平面上至少有一直线过A而与a不相交的结论可利用其他公理推证出来，故不必列为公理。）

Ⅴ．**连续公理**

Ⅴ$_1$（Archimedes公理）：若AB与CD为任意二截段，则在AB直线上存在一组点A_1，A_2，…，A_n使得AA_1，A_1A_2，…，$A_{n-1}A_n$都合同于CD而使B位于A和A_n之间。

（注意．实际可使B位于A_{n-1}和A_n之间。又因CD是任意的，且$A_{n-1}A_n \equiv CD$，故即表明A_{n-1}与A_n可与B任意靠近，由此推论，存在点列$\{A_n\}$使得$A_n \to B$。）

Ⅴ$_2$（线性完备性公理）：凡满足公理Ⅰ$_1$，Ⅰ$_2$，Ⅱ，Ⅲ和Ⅴ$_1$的直线上的一切点构成的点集不可能再扩大。

（注意．这条公理保证了直线上的一切点可以和实数一一对应起来，有时称此为康托公理。）

希尔伯特曾根据上述五组公理导出了欧氏几何的若干基本定理。后来人们继续演绎推导，直至证得欧氏《几何原本》的所有定理，并且充分显示出希尔伯特关于欧几里得几何公理系统的陈述已经免除了《几何原本》的不足之处。

§5　关于公理系统的相容性问题

通常把由一组原始概念和公理刻划的数学理论称为一个数学系统。而一个数学系统的相容性问题就是指刻划它的那个公理系统的相容性问题。

关于相容性证明这一概念的产生和历史发展的背景是这样的，自从罗巴切夫斯基几何诞生后，由于罗氏平行公理是如此地为常识所不容，这才激起了人们对于数学系统的无矛盾性证明的兴趣和重视。虽然在罗氏公理系统的展开中一直没有出现矛盾，却不能保证它在今后的展开中一定不出矛盾。后来，庞卡莱在欧氏系统中构造了一个罗氏几何的模型，亦即在欧氏平面上划一条直线a而使之分为上、下两个半平面，把不包括这条直线在内的上半平面作为罗氏平面，其上的欧氏点当作罗氏几何的点，把以该直线上任一点为中心，任意长为半径所作出之半圆周算作是罗氏几何的直线，然后对如此规定了的罗氏几何元素一一验证罗氏几何诸公理全部成立。在这里，我们朴素地来说一说罗氏平行公理是成立的。如下图所示，过罗氏平面上任一罗氏直线l外的一

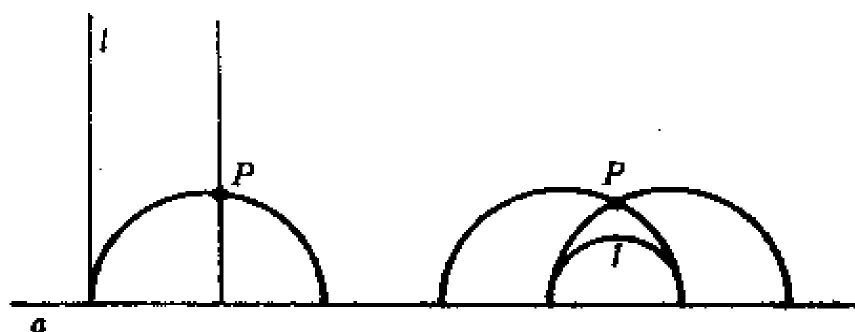


图4-3

点P，确实可以作出两条罗氏直线与l平行。这里要注意的一点是欧氏直线a上的点不是罗氏系统的几何元素，故两个半圆相交于直线a上某一点则视为相交于无穷远点，从而在有穷范围内永不

相交。这样一来，如果罗氏系统在今后的展开中出现了正、反两个互相矛盾的命题的话，则只要按如上规定之几何元素间的对应名称进行翻译，立即成为互相矛盾的两个欧氏几何定理。从而欧氏系统就矛盾了。因此，只要承认欧氏系统是无矛盾的，那么罗氏系统一定也是相容的。这就把罗氏系统的相容性证明通过上述庞卡莱模型化归为欧氏系统的相容性证明。这种把一个公理系统的相容性证明化归为另一个看上去比较可靠的公理系统的相容性证明，或者说依靠某一个数学系统的无矛盾性来保证另一个数学系统的协调性叫做数学系统的相对相容性证明。但是，人们本来对于欧氏系统的相容性没有怀疑过，却因罗氏系统的相容性要有欧氏系统的相容性来保证，从而导致对欧氏系统相容性的重重疑虑。人们还在罗氏系统的展开中发现，罗氏几何空间中的极限球面上也可构造欧氏模型，亦即欧氏几何的全部公理能在罗氏的极限球上实现，这样欧氏几何的相容性又可由罗氏几何的相容性来保证。这说明欧氏与罗氏的公理系统虽然不同，但却是相对相容或互为相容的。人们当然不满足于两者互相之间的相对相容性证明，因为看上去较为合理的欧氏系统的无矛盾性竟要由看上去很不合理的罗氏系统来保证，这是难以令人满意的。因此，必须重新寻求欧氏系统的相容性证明。由于那时已经有了解析几何，等于在实数系统中构造了一个欧氏几何的模型。这就把罗氏系统的相容性进一步归结到了实数论的相容性。但实数论的相容性如何？这样的归结和提问永远不会完结。Dedekind把实数论的无矛盾性归结到了自然数系统的无矛盾性，而Frege又把自然数系统的相容性归结为集合论的无矛盾性。然而，集合论的无矛盾性又如何？至今还是个谜，以致公理系统的这种相对相容性证明至今还是一场空。希尔伯特早就提出：不能依靠相对相容性证明来解决问题，而应该搞直接的相容性证明，这将在第9讲中作进一步讨论。

固然希尔伯特的几何公理系统从纯逻辑观点看尚未彻底解决

协调性问题，但只要明确引入自然数无矛盾的基本假设作为公设之后，该公理系统在相对意义下的无矛盾性就获得保证了。

人们在理性思维上总是习惯于希望通过逻辑推理证明一切。岂知某些具有“无限性”飞跃结构的概念系统往往越出有限步逻辑推理判断的范围之外。因此，如果懂得点概念思维的辩证法，也就能够较自觉地去识别并避免徒劳无功的尝试了。

最后，值得说明一下，正因为希尔伯特几何公理系统中的点、线、面、位于、通过等名词都无非是一批抽象元素及其关系的代名词，因此对它们可以赋于各种各样的具体解释。如果把它们解释作古典欧氏几何（平面几何与立体几何）中的对象，则得到二维及三维欧氏几何。特别，如果我们把公理中的点与直线分别反过来解释成普通欧氏几何中的直线与点，便可得出原定理的对偶定理。正因为公理中的点与线皆为抽象元素，故在名词上互易其位也无不可。所以，就有一般形式的**对偶原理**：在任何一条涉及点、线关系的平面几何定理里，如将点、线位置互换，则所得命题仍成立（后一命题即称为前一定理的对偶命题）。

上述对偶原理很有用，它可以帮助我们在几何证题上一举两得。例如，当我们证明了著名的巴斯卡（Pascal）圆内接六边形定理后，也就立即可得布列安训（Brianchon）的圆外切六边形定理。

从对偶原理的导出，已使我们看出抽象化的公理系统确实概括了较丰富的内容。

关于公理系统的独立性要求，我们仍以欧氏与罗氏两个几何公理系统为例，在相对相容意义下略加讨论。如前所述，在欧氏与罗氏两个几何公理系统中除了欧氏平行公设与罗氏平行公理互为相反之外，其余的公设、公理和原始概念均相同。通常人们把两个公理系统的公共部分称为绝对几何公理系统。因之，欧氏平行公设在欧氏几何公理系统中是否独立于其他公理一事，无非就是欧氏平行公设能否在绝对几何公理系统中作为定理而证明之。

而只要罗氏几何公理系统是无矛盾的，就确保了欧氏平行公设对于绝对几何公理系统的独立性。否则，若能在绝对几何公理系统中把欧氏平行公设作为定理来证明的话，则罗氏几何公理系统便是矛盾系统了，因为此时欧氏平行公设和它的否命题（罗氏平行公理）在系统中同时成立。完全类似地，由欧氏几何公理系统的无矛盾性也能确保罗氏平行公理对于系统中其他公理在逻辑上的独立无关性。因此，一般地说来，要证明某一公理系统Σ中某一公理对系统中其它公理在逻辑上的独立性，只要构造并证明公理系统$\Sigma' = (\Sigma - A) + \neg A$是无矛盾的即可。

关于公理系统的完备性要求，不妨参考H. B. Ефимов著，裘光明译，《高等几何》上册，p.223—226。现将对于希尔伯特公理系统的完备性的描述和理解介绍如下。

先让我们给出关于同一个公理系统的不同模型之间的同构概念。假设Σ为某一个已知的公理系统，而Σ在两个不同的对象集合S和S'上分别构造了两个模型Σ_S和$\Sigma_{S'}$。如果在这两个模型的对象中间可以建立这样的一一对应，使得对应元素有同样的相互关系，则称Σ_S和$\Sigma_{S'}$同构。若以几何模型作解说时，所谓有同样的相互关系，就是指当Σ_S的点A和直线a对应于$\Sigma_{S'}$的点A'和直线a'时，如在Σ_S中A落在a上，则在$\Sigma_{S'}$中必有A'落在a'上，如此等等。

若把希尔伯特关于欧几里得几何公理系统中第一组的前三条公理$\text{I}_{1,2,3}$作为一个独立的公理系统，记为Σ。把某一三角形的三个顶点叫做点，三条边叫做直线。这样，我们就有了一个共有六个对象作为元素的集合σ，这个三角形正是Σ在σ上的一个模型，记为M。为什么？因为容易验证I_{1-3}三条公理的要求在这里得到满足。例如公理I_1要求过任何两点A和B有一条直线a，此处对于三角形的任何两个顶点都有一条边连接它们；其余类此验证。

我们在Σ中加入新的公理，扩大到I_{1-8}，把它作为一个独立的公理系统，记为Σ_1。今把一个四面体的四个顶点叫做点，六条

棱叫做直线，四个面叫做平面，如此得到四个点、六条直线、四个平面共有14个对象作为元素的集合σ_1。容易验证，所给四面体这个模型M_1可使公理Ⅰ$_{1-3}$一一得到满足。

再对Σ_1加入新的公理，直至扩大为欧氏几何全部公理所构成的系统Σ_2，我们可以通过解析几何在实数集σ_2上构造出欧氏几何的模型M_2。

须知M_1和M_2也可视为Σ在σ_1和σ_2上的模型。因为Σ的几条公理Ⅰ$_{1-3}$显然在M_1和M_2上得到满足；至于同时还有别的公理也成立这一事实，我们可以不问。于是，此处所给出的Σ的三个模型M、M_1和M_2显然都是互不同构的。因为在这三个模型之中的任何两个，连一般的一一对应也不存在，更谈不上在对应的元素间保有相同的关系了。

从上面的讨论中可以看到，一个公理系统中的公理愈少，则选取它的模型的自由度就愈大，就是说，当我们不断地把一个公理系统扩大（当然要求加入的新公理对于原有的公理来说保有独立性和矛盾性）的时候，则能成为公理系统的模型的种类就越来越少。基于这样的讨论，我们可以把一个公理系统的完备性概念确切地叙述为：如果已知的公理系统的所有的模型都是互相同构的，则该系统称为完备的。

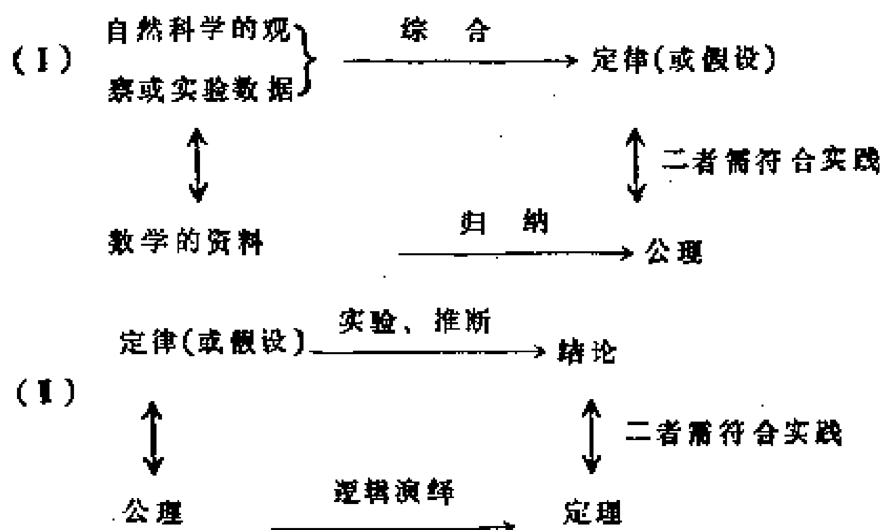
基于以上的完备性概念的讨论，可以验证希尔伯特所给出的欧几里得几何公理系统是一个完备的公理系统，因为它的任何一个模型都与该公理系统的笛卡尔模型（即通过解析几何在实数域上构造出来的模型）是同构的，因之互相之间也是同构的。

§6　略谈自然科学中的公理化方法

下面摘引袁相碗同志一文〔载《南京大学学报》（自然科学版）1980年第二期〕中的一段分析以供参考

在自然科学的具体研究工作中，公理方法，特别是它的逻辑

思维有着重要的影响。这可由下图表示：



上述（Ⅰ）是由“果”到“因”。这对自然科学是重要的。例如，天文学家第谷（Tycho）勤于观察，得到了有关五大行星运行的大量数据，但他缺乏公理推导的训练，对所得的大量观察材料无法进行综合分析。而刻普勒（Kepler）运用数学的公理推导的方法，经过精心计算，归结出著名的行星三大运动定律。后来，牛顿更进一步运用公理推导的方法，提出“万有引力论”，由这个学说遂一导出了刻普勒的三大定律，从而使天文学和物理学又前进了一大步。试想，如果没有公理方法，把杂乱无章的数据追溯到几条简单的基本规律，而单凭肉眼观察，谁看见过行星的运行轨道？谁看见过行星运行过程中所扫射的面积？谁又看见过两物体之间互相吸引的引力呢？再者，如果没有公理推导能力，谁能断言第谷所积累的观测数据恰恰可用刻普勒三大定律来说明？而三大定律又恰恰可由牛顿的万有引力定律来导出？

（Ⅱ）是由“因”到“果”。这在自然科学的研究中也是常见的。在自然科学史上，往往采用公理方法，从基本假说或少数定律出发，进行理论推导，看看会推出哪些尚未观测到的或尚未发现过的现象，然后再用实验去验证。例如，人们曾根据爱因斯坦的广义相对论，推测有“黑洞”存在和引力波存在。人们在探

索过程中，根据这些解释了许多天文学上原先无法解释的现象。可见，应用公理方法还可做出科学预见并澄清疑难。但是必须强调，在自然科学研究中应用公理方法时必须和实验观察互相配合，互相参照，相互促进。否则就很少成效，甚至会导致荒谬。

最后，我们再按照反映论的观点为实体公理化的作用补充说几句。一般说来，由于形式公理化较之实体公理化毕竟是一种具有更高层次的科学抽象形式，一方面这种抽象形式能够更深刻地更突出地反映事物的某些本质，另一方面这种抽象过程又必然扬弃掉事物客体的种种次要环节，因此最后反倒不能较细致地逼真地描绘出事物内在本质中相互联结在一起的诸环节 就这一点来说，实体公理化却比形式公理化更加贴近实体对象的本性和体貌。所以在自然科学领域中，如果片面地追求纯粹的形式公理化而放弃实体公理化，则对表述和反映科学真理内容反倒会形成片面性，因而未必是很明智的做法。其实，即使对于包含着种种无限过程的数学理论体系来说，要按希尔伯特证明论的方案来实现彻底的形式公理化，也已包含着不可克服的困难，而改用实体公理化，反而更有利于反映一个分支中的数学科学的真理全貌。总之，实体公理化仍不失为一种有用的科学分析方法，它的功效和作用是并不能由形式公理化方法所替代的。

第5讲　关于数学的结构主义

§1　结构主义学派的形成过程

结构主义学派又称**布巴基学派**，它是本世纪三十年代以后开始形成的一个数学学派．布巴基（Bourbaki）并无其人，只是一个假名．实际上它是这一派人物著书立说时共同采用的一个笔名．这一派的首要人物包括J.Dieudonne，A.Weil，C.Chevalley，H.Cartan．他们有许多追随者．在四十至五十年代曾盛极一时．

大约在三十年代左右，法国出现了一批年青有为的数学家，他们不满意于老一代数学家的守旧作风，要求跳出老一套的圈子，希望闯出一条新路来，使数学能重新获得生气勃勃的发展．他们建立了讨论班，共同合作研究．他们运用了公理化方法，试图把一些数学分支中进行论证的最基本、最重要的出发点分离出来，并加以比较．于是便形成了**结构**的概念．

这一学派的主张所以称为结构主义，是因为他们认为数学各分支应按**结构性质**来划分，应用公理方法按**结构观点**来重新整理各个数学分支．他们希望全部数学或大部分数学都能纳入各种结构系统中去．他们的雄心很大，计划很宏伟．

他们分工合作出版了一大套书籍，名为《数学原本》（Elements of Mathematics），作者署名Nocolas Bourbaki，1939年出版了第一卷，到1971年共出版三十六卷，仍未宣布写完．在《数学原本》内，大部分数学分支迄今都经过仔细的结构分析，并安放到适当的位置上．

特别有趣的一本书名为《数学史原本》，于1969年出了第二

版。书中全按结构观点叙述了数学发展史。

§2 布巴基学派的一般观点

仔细说来，布巴基学派的基本观点是认为，全部数学或大部分数学都可以依照结构的不同而加以分类。用公理化方法抽象出各个学科的各种结构，找出各数学分支间的结构差异，这样，就会获得各数学分支的**内在联系的清晰图景**。他们还认为数学的发展，无非是各种结构的建成和发展而已。

在谈到数学发展的过程时，他们作了一个通俗的比方，他们说："数学好比一座大城市。城市中心有些巨大的建筑物，就好比是一个个已经建成的数学理论体系。城市的郊区正在不断地并且多少有点杂乱无章地向外伸展，他们就好象是一些尚未发育成型的正在成长着的数学新分支。与此同时，市中心又在时时重建，每次都是根据构思更加清晰的计划和更加合理的布局，在拆毁掉旧的迷宫似的断街小巷的同时，将修筑起新的更直、更宽、更加方便的林荫大道通向四方，……。"

由上所述，可见结构主义本质上可以看成是近代形式公理化思想的一个发展。数学公理化思想着重于探讨每门数学的公理化方法，而结构主义则采取全局观点，着重分析各个数学分支之间的结构差异和内在联系。而且对每门数学也着重分析其结构特征，或关于它的某些基本结构的组成方式。至于建成每一数学分支的具体方法也还是公理化方法。

§3 数学结构的分类

先说一说**结构**的概念。一个抽象的集合不过是一组元素而已，无所谓结构。但引进了运算或变换，就形成了结构。结构中必须包含着元素间的关系，这些关系通常是由运算或变换联系着

的。

例一 最早提出的一种简单结构就是群。群是很普遍的一类代数结构。例如，伽罗瓦（Galois）群的概念已超出了数的运算范围。凯雷（Cayley）继伽罗瓦之后，提出了抽象群的概念，其内容对象更为丰富。举例言之：

正负整数添上零按加法构成群——{**N**；+}。

正有理数按乘法构成群——{**R**，×}。

几何中有种种变换群，如射影变换群。

晶体分子排列中有置换群。

旋转运动中有转动群。

分析数学中的许多函数类也构成群，如多项式按加法构成群。

代数中的向量按加法也构成群，等等。采用结构观点考察群时，不是注意各种具体的数学集合，而是注意集合（按群的运算）所表现的内在关系结构，如子群间的关系结构等。为此，就必须注意群运算所具备的基本性质，比如群的公理。这些便是我们所要研究的对象。

众所周知，群结构观点已渗透到一切数学部门中。

例二 几何学可以按照不同的群结构来分类。这是德国数学家克莱因（Klein）于1872年在德国爱尔朗根讲演中首先提出的思想方法。数学史上称它为爱尔朗根（Erlangen）纲领。

按克莱因认为欧氏几何的对象就是研究**刚体变换群**作用下的不变性质（如长度不变，角度不变等）。射影几何研究的是**射影变换**下的不变性质。拓扑学研究的是**拓扑变换**（即一对一且具有双方连续的变换）下的不变性质（俗名橡皮体上的几何学）。如此等等。

布巴基学派将数学结构分为三大类：

（一）**代数结构**。由离散性对象加运算构成的结构系统，如群、环、域、代数系统、范畴、线性空间等。

（二）**序结构**。如半序集、全序集、良序集等。

（三）**拓扑结构**。如拓扑空间、紧致集、列紧空间、连通集、连续性及完备性空间等。

上述三种结构叫做**母结构**。由此可导出各种子结构。还可有各种交叉，形成“分支结构”。举例如下：

例一 拓扑群是群结构上再定义拓扑结构的一门学科。

例二 H空间（希尔伯特空间）是线性空间（代数结构）添上内积型拓扑（拓扑结构）所构成的数学系统。

例三 B型空间（巴拿赫空间）即完全（完备）、赋范、线性空间。那也是一种交叉而成的分支结构。

数学中有些对象（如数直线）看来很简单，其实，如果进行结构分析，即可发现其结构相当复杂。

§4 数直线结构分析

所谓数直线**R**，就是由全体实数构成的一维欧氏空间。现在我们来解剖**R**的结构，可以发现**R**是一个**完备的**Archimedes**全序域**，它乃是由代数结构（域）、序结构（全序）、拓扑结构（完备性结构）形成的分支结构。

简记**R**＝{实数；＋，×，≤}。容易验明：

（1）**R**具有**环结构**。对＋、×各个运算都满足交换律与结合律，0是加法单位元，1是乘法单位元。对＋、×还满足分配律。故对两种运算而言构成一个环，而且对乘法还有逆运算，故**R**实际上还构成一个域。

（2）**R**具有**序结构**。对序关系≤满足三条基本性质，即传递性，对称性与可比性。故**R**是一个全序结构。此外，对加、乘（＋，×）而言还满足保序律。这种保序性使代数结构与序结构具有协调性，因而合在一起能做成新的结构。

（3）**R**具有**连续性结构**。Archimedes公理显然是满足的

（即任给 $x>0$，$y\geqslant0$，总存在自然数n致使$y\leqslant nx$）。

对上述有序域R中的元可定义绝对值，从而引出距离概念和邻域概念。由此可得到极限概念和基本序列概念，且甚易验明 R 具有完备性，即凡基本序列的极限值也都含于R 内。故结论是：R 是一个**完备的亚几米德有序域**。它是代数结构、序结构和拓朴结构三者联合而成的具有协调性的一种交叉结构。

注意，(1) 在R上可以展开全部数学分析；因为在其上极限过程可以畅行无阻，凡所得极限恒在R之中．(2) 完备序域R的子域——有理数域Q也是一个亚几米德有序域，但不完备，因为收敛的有理数列的极限未必属于R．因此在有理数域上无法建立分析数学．(3) 确实存在着非标准数域的例子．例如利用"超幂"（Ultrapower）构成的非标准实数域R^*就是一个非亚几米德的完备有序域，在这域中存在无穷多个非标准无穷小实数．例如，设α为无穷小，则$\alpha\in R^*$．于是对任给标准实数$y>0$，却不可能找到自然数n致使$n\alpha>y$，原因是$n\alpha$仍为无穷小

§5　略谈拓扑结构

一般都比较熟悉代数结构，而序结构概念也比较简单，所以这里专就拓扑结构略作讨论。

所谓拓扑结构，就是能够描述极限的那种结构。要描述极限，就要有"靠近"或"邻近"这样一些概念，这就需要距离概念。因此对于一个集合中的元素间只要能引进"距离"（或度量）的定义，即可形成一个拓扑结构、所谓距离拓扑。我们知道，近代泛函分析中讨论的各种线性赋范空间，其范数(或模数)概念就是一种广义绝对值概念，由此也可导出距离概念。所以凡赋范空间都具有距离拓扑结构。

当然，也可直接用邻域（neighborhood）这种概念来界定极限概念。例如，在欧氏空间中一维邻域便是开区间，二维邻域可以是开圆域，三维邻域可以是开球域，等等。

为了在最一般形式下引进拓扑概念，最简便而且比较符合直观的办法就是先引进一般意义下的邻域和邻域系的概念。如同确

立别的数学分支或数学体系的公理那样，必须对邻域需要具备的基本性质作一全面分析．然后利用反映基本性质的条款来界定邻域概念．反映基本性质的条款也就是**领域公理**．当然，这些公理必须是无矛盾的，而且是相互独立的．它们还必须是完备的，即确实足以刻划所需要的极限概念．

现代的拓扑学教程中，都要谈到邻域公理，这组公理可叙述如下：

设在一个抽象点集E中，分离出一些子集称之为邻域．若U是邻域，$x\in U$，则也称U为x的邻域．邻域系$\{U\}$须满足下列四条公理：

（1）E中每一点至少有一邻域．

（2）一个点x的两个邻域U_x与V_x之交$U_x\cap V_x$仍为一邻域．

（3）若N包含U_x，即$E\supset N\supset U_x$，则N也是点x的一个邻域．

（4）在点x的任一邻域V_x中必含有x的邻域U_x，使得当$y\in U_x$时总有$U_y\subset U_x$．

上述公理中的（1）、（2）、（3）陈述较简单，其重要性在直观上也比较明显．看来只有公理（4）似乎略复杂些．事实上，人们在作分析数学领域中的命题论证时，时常需要用到这条公理作为分析推理的依据．这条公理的意思无非是肯定每一点的某个邻域中的一切点也都有邻域．特别，每一点的开区域都具备这个性质．在分析学中联系（4）和（2）往往能保证选取一串越来越小的邻域使之以点x为极限点．由此公理（4）的重要性可见一斑．

由邻域系的概念就可以导致极限概念．凡具有邻域系的集合就称为**拓扑空间**，邻域系称为**拓扑结构**．正如本讲§3中提及的，所谓拓扑学就是研究拓扑变换下不变性质的一门几何学．我们知道，关于空间形式或几何形态的拓扑变换（即原象与映象双方连续的一一映射）是一种广泛地存在于现实世界的变换方式，所以

拓扑学研究对象的重要性也就不言而喻。试看：河边柳树枝的飘摇，河水的流动，还有绳子的打结、工厂锻工车间的“压模”等都涉及几何形体的变化方式，哪个例子不可纳入拓扑变换的范畴中去呢？如所知，不只是理论数学各分支，而且连现代的生物学与经济科学领域也都借助于拓扑学工具，其原因就是因为拓扑学的研究对象和内容太广泛太丰富了，换言之，在现实世界中具有**拓扑结构**的原型太常见了。

布巴基学派非常重视代数结构。同时他们又把拓扑结构列为母结构之一，这自然是理所当然的，也是十分必要的。

§6　略谈同构概念

从结构的观点出发来分析问题，同构的概念是一个非常重要的概念。这是因为凡具有同构性质的一些结构，在本质上都可看成是同一种结构；在研究问题时当然只须抓住一种结构进行分析即可，而无需浪费重复性的劳动。

假定有两个集合M和N，他们对分别定义的运算$\circ$和$*$各自满足同样一组公理。如果存在一个一一对应法则φ，使M和N中的元以及各自的运算结果都能一一对应起来，如：$x\in M$，$x\longleftrightarrow y=\varphi(x)\in N$，且

$$x_1,\ x_2\in M,\ x_1\circ x_2\in M, x_i\longleftrightarrow y_i=\varphi(x_i)\in N, (i=1,2),$$

$$x_1\circ x_2\longleftrightarrow y_1 * y_2=\varphi(x_1\circ x_2)\in N.$$

这样，就称$\{M;\ \circ\}$与$\{N;\ *\}$是一对同构系统。

当然，系统间的同构概念也可推广到具有多重运算的情形。

数学主要是研究结构，至于结构采取什么形式自然无关紧要。因此，在一类同构的系统中只须选定一个结构作为研究对象即够。

下面我们举复数域的四种不同表现形式作为同构系统的简单例子。

令a，b表实数，$i=\sqrt{-1}$为虚数单位，则按如下的各种一一

对应关系可知复数$z=a+bi$有四种表现形式：

$$a+bi \longleftrightarrow (a,b) \longleftrightarrow \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \longleftrightarrow \overrightarrow{OA}$$

对于坐标表示法$(a,\ b)$而言，可规定加法和乘法运算如下：

$$(a,\ b)\ +\ (c,\ d)\ =\ (a+c,\ b+d),$$

$$(a,\ b)\ \cdot\ (c,\ d)\ =\ (ac-bd,\ ad+bc)\ .$$

显然这与$a+bi$，$c+di$间的运算结果是对应一致的。

又对于矩阵表示法而言，如果

$$a+bi \longleftrightarrow \begin{pmatrix} a & b \\ -b & a \end{pmatrix},\quad c+di \longleftrightarrow \begin{pmatrix} c & d \\ -d & c \end{pmatrix},$$

则易见

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}+\begin{pmatrix} c & d \\ -d & c \end{pmatrix}=\begin{pmatrix} a+c & b+d \\ -(b+d) & a+c \end{pmatrix} \longleftrightarrow (a+c)+(b+d)i,$$

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}\cdot\begin{pmatrix} c & d \\ -d & c \end{pmatrix}=\begin{pmatrix} ac-bd & ad+bc \\ -(ad+bc) & ac-bd \end{pmatrix} \longleftrightarrow (a+bi)(c+di).$$

这些都反映了复数间相同的运算规则。同样，对复数除法（乘法逆运算）也是完全对应一致的。因此，无论是由坐标表示法或者矩阵表示法，由相应运算规则确定的数学系统都是互相同构的，它们在本质上都代表同一个复数域。细节的验证读者容易自行补出，兹从略。

进一步，在上述复数域上还可引进拓扑结构。这只须引进范数（绝对值）概念即可：

$$|a+bi| = |(a,b)| = |\overrightarrow{OA}| = \left|\begin{pmatrix} a & b \\ -b & a \end{pmatrix}\right| = \sqrt{a^2+b^2}.$$

这样它们构成的系统都成为同一种半序域p，其中序的关系可按下列方式规定：

$$(a+bi) \leqslant (c+di) \Longleftrightarrow a \leqslant c \quad 且 b \leqslant d.$$

显然$3+4i$与$4+3i$之间就没有序的关系。

如果对每一复数$z_0=x_0+y_0i$，把满足条件

$$|z-z_0| = |(x-x_0)+(y-y_0)i| < \delta,\quad (\delta > 0)$$

的点集$\{z\}$称为z_0点的δ邻域，则所有这种形式的邻域（其中δ为任意正数）及其拓扑变换后的一切子集显然满足“邻域系公理”（1）——（4）。因此范数的引入也就产生了拓扑结构。

§7　略评结构主义

确实，数学的各分支可以按照它们的结构性质加以区别和归类，所以布巴基派的基本观点是正确的。他们试图把大部分数学都纳入各种结构系统中去，并通过比较使得各分支的内在联系与区别展示出一幅清晰的图景，这对于全面地整理数学来说，无疑是十分合理的作法。

但是这一派的方法论，专注重于数学形式结构特征的分析与比较，可以说是一种关于已经形成了的数学部门的**回顾性的逻辑分析**，而不是展望和探索新领域的方法。换言之，结构主义并不注意研究如何从现实世界中提取新的数学模型，开辟新的数学园地。其主要兴趣不过是整理和分析。因此，结构主义的基本思想方法看来并不是一种发明创造的方法。

结构主义的观点反映在数学教学领域，曾经导致现代热闹一时的新数运动。新数在以往十多年里曾在欧美诸国进行过多次试验，看来并不成功。按照新数的教学观点，就是要给学生一开始就讲授最一般的数学结构系统，于是象集合论与抽象代数乃至数理逻辑就成为必要的基础工具。青年学生的记忆力和模仿性都较强，当然开始时也容易照样画瓢地在数学形式演绎系统中表现一番。但毕竟由于缺乏生动的直观背景，违反一般人的正常的认识过程，所以新数在数学教学中的失败，乃是理所当然之事！当然也有极少数的早熟儿童或非常杰出的青少年能从新数中得到好处，但那只是极个别的例外。

由上所述，可知按结构主义的纯演绎形式讲授数学题材的观点即新数的观点。这观点表现在实践上的最大失败就是无法激发

出青年学生们的直觉想象能力。可是直觉与想象正是数学创造力的源泉。如此说来，在数学教育或教学领域中，结构主义的表述方式看来是没有多大价值的。事实上，布巴基学派提供的《数学原本》数十卷，远远不能同历史上欧几里得《几何原本》在教育上的作用相比；它们在现代数学教育上的影响甚微。当然，对专业研究工作者还是有一定参考价值的。

第6讲　代数方程根式解法与伽罗瓦的群论思想方法

§1　代数基本定理与根式解法研究简史

代数方程根式解法的研究有很悠久的历史。大家知道，一个实系数的代数多项式在实数域中只要能分解成一些一次因式与二次因式的乘积，则方程的各个根便可立即得出。一般说来，二次因式为零所表现的二次方程会给出一对共轭复数根。

是不是每个n次实系数多项式都能有如上所述的因式分解形式呢？当初人们对此是有怀疑的。例如莱布尼兹（Leibniz，1646—1716）就不相信这件事。对此问题首先得出正确答案者是欧拉。但他只对次数$n \leqslant 6$的代数多项式给出了证明。

代数方程式论的一个中心问题是：每一个实系数多项式方程是否总有一个实根或复根？这就是代数方程根的存在性问题。显而易见，只要能证明凡n次方程必有一根，则根据余式定理和归纳法便可推出n次方程必有n个根（包括重根）。

对方程根存在性命题首先试图证明的是达朗贝尔（Dalembert），接着欧拉也给了一个证明。但他们的证明都是有缺陷的。拉格朗日（Lagrange，1736—1813）于1772年又重新证明了根的存在性命题，后经高斯分析，发现其证法中曾把实数的各种性质应用于尚未证明其存在的根上，所以该证明仍然是很不严格的。高斯于1799年在哥丁根（Göttingen）大学所完成的著名的博士论文即以方程根的存在命题为主题，在该文中才首次严密地证明了上述代数学基本定理。

高斯证法的基本思想如下。设$f(z)$为n次实系数多项式，记

$z=x+iy$，（$i=\sqrt{-1}$，x，y为实数），考虑方程$f(x+iy)=u(x, y)+iv(x, y)=0$，即$u(x, y)=0$，$v(x, y)=0$。这里$u(x, y)=0$与$v(x, y)=0$分别表示$OXY$坐标平面上的两条曲线$C_1$与$C_2$。于是通过对曲线的定性分析，可以证明曲线之一（例如C_1）必有一段连续弧联接着另一曲线（如C_2）所划分的两个区域中的两个点，所以必有一个交点$z^*=a+ib$。从而得出$u(a, b)=v(a, b)=0$，也即$f(a+ib)=0$。因此z^*便是代数方程$f(z)=0$的一个根（这个证法的思想方法十分自然，欲知其详，请参考美国Birkhoff－Maclane所著《现代高等代数》）。

高斯对上述代数学基本定理十分重视，一生中曾给出了四个证明。1816年与1825年给出的第二、第三个证法基本上用到了柯西积分定理的思想。高斯七十一岁高龄时在Göttingen科学杂志上还公布了第四个证法。在这最后一个证明里多项式的系数可以是复数系数。

复变函数论发展后，代数学基本定理已经作为更一般定理的推论。例如，熟知的刘维尔定理断言：凡在复数平面上处处解析的函数如果有界，则必为常数。换言之，凡不等于常数的整函数必无界。今假定n次代数方程式$f(z)=0$（$n\geqslant 1$）在复数平面上无根存在，即对一切复数z恒得$f(z)\neq 0$。显然$f(z)$不是一常数且于$|z|\to+\infty$时，$|f(z)|\to+\infty$。于是对给定正数M必有R，使得于$|z|>R$时恒有$|f(z)|>M$。记

$$\inf_{|z|\leqslant R}|f(z)|=d>0.$$

则由假定知$g(z)=1/f(z)$为在复数平面上处处解析的函数。但

$$|g(z)|=|1/f(z)|<\frac{1}{M}\quad(|z|>R);$$

$$|g(z)|\leqslant\frac{1}{d}\quad(|z|\leqslant R).$$

故推知$g(z)$为有界，据刘维尔定理应得$g(z)=$常数，也

即$f(z)$ = 常数。这与$f(z)$本性矛盾。因此$f(z)=0$在复数平面上必有根存在。可见代数基本定理乃是刘维尔定理的简单推论。

数学史上常有这种现象，用后来发展了的工具去处理老问题时，常能获得最简捷的解决方法。

当然，代数学基本定理只是从原则上肯定了代数方程根的存在性，至于方程的根如何根据方程的系数去计算的问题那是另外的问题。特别，可否利用系数间的四则运算与开方根运算表出方程根的问题，那是更难的问题。这后一问题叫做关于代数方程的根式求解问题。这个问题的提出是很自然的，因为当人们发现了二次方程、三次方程与四次方程的一般求根公式之后，也就必然会提出五次或五次以上的代数方程有无一般求根公式的问题。

最初，莱布尼兹曾研究了五次方程的根式求解问题，但未获成功。后来高斯专门研究了二项方程$z^n-1=0$的根式求解问题，取得了重要成果。特别，如果n是质数且能表成$n=2^{2^k}+1$的形式时，则$z^n-1=0$的各个根还可利用有理数的有限次五则运算+、-、×、÷、$\sqrt{\ }$（开平方）表示出来，这个结果在尺规作图中的分圆问题上有重要应用。

继拉格朗日与高斯之后，鲁飞尼（Ruffini）、阿贝尔与雅可比曾各自探索了一般五次方程的根式求解问题。阿贝尔与雅可比曾一度认为“已解决了问题”，但很快就发现其作法是不对的。鲁飞尼首先按照正确的思路“证明”了一般五次方程根式解法的不可能性。但他的证法中用到了一个关键性命题而未加证明。所以鲁飞尼的工作方向虽然是正确的，但提供的证明却是不完整的。

阿贝尔在高中读书时就阅读拉格朗日、高斯有关方程式论的著作。开始时他利用高斯处理二项式方程的手段去研究五次方程，曾一度以为能用根式解出五次方程，在很快发现了错误之后，他便敏感地猜想到一般五次方程不可能用根式求解的结论。接着，

他成功地证明了一条定理，今称阿贝尔定理，那实际就是鲁飞尼论证中所忽略的一个关键性的基本命题。由此基本命题阿贝尔便证明了："高于四次的一般代数方程不可能有一般形式的根式解。"这是数学史上的一项重要成就。

由于阿贝尔并未见到鲁飞尼的工作，他在1826年完成的证法中不免有许多与鲁飞尼重复的部分，所以文章写得不必要的复杂。他的论文中还有一点小错误，即关于函数分类的错误。幸好这点错误无关紧要。后来在他1829年去世前，还重写了两个更精细的证明，寄给了勒让特尔（Legendre）与克雷耳（Crelle）并发表于克雷耳杂志。1879年恰好过了半个世纪德国数学家克朗内克（Kronecker）在阿贝尔的思想基础上又给出了一个简明、直捷、严格的证明。

固然高于四次的一般代数方程（带有文字系数者）不可能有一般形式的根式解，但有些特殊类型的方程（例如二项方程）却仍然可以得到根式解。因此，高次代数方程何时存在根式解的问题——即代数方程的可解性问题乃是一个需要进一步解决的问题。

对上述问题的研究，拉格朗日曾迈出了重要的一步。他用某种统一的观点分析了二次、三次、四次方程的根式解法之后，曾预见到一般方程的可解性问题最后都将归结到关于诸根的某种排列置换的问题，而且他还设想了一种理论上的（或原则性的）利用根式求解方程的步骤。不幸的是，拉格朗日按照自己的一般设想方案去处理五次方程时碰了壁，因为事情竟然复杂到了使他无法进行下去的地步，于是只好放弃了上述工作计划。

可是拉格朗日的洞察力却启发了年轻的阿贝尔与伽罗瓦（Galois，1811-1832），他们在继承了拉格朗日留下的宝贵遗产的基础上，各自作出了重大的贡献。尤其是，犹如划破黑夜长空的一颗彗星——伽罗瓦的出现，开创了置换群论的研究，从而确立代数方程的可解性理论，彻底解决了一般方程的根式解难

题。有关他们的基本思想方法将在后面各节中加以论述。

§2 拉格朗日的思想方法与阿贝尔定理

人们在16～17世纪时已熟知代数方程根与系数的关系。设x_1，x_2，…，x_n为下列n次方程

$$f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_n=0 \qquad (a_0\neq 0)$$

的n个根（其中容许有重根）。则

$$f(x)=(x-x_1)(x-x_2)\cdots(x-x_n)=0$$

如果将上式乘积展开出来，则比较同次幂系数可得联立方程

$$\begin{cases}\sum_{k=1}^{n}x_k=-a_1/a_0\\ \sum_{i<k}x_ix_k=a_2/a_0\\ \cdots\cdots\cdots\cdots\\ x_1x_2\cdots x_n=(-1)^na_n/a_0.\end{cases}$$

这就是根与系数的韦达关系，其中各等式左端均为诸根的对称式，即诸根互换位置后等式形状均不变。

所谓代数方程$f(x)=0$的根式解，实质上可以看作是上述联立方程的代数解，即利用文字系数a_0，a_1，…，a_n间的有限次四则运算与开方根运算$\sqrt{\ }$，$\sqrt[3]{\ }$，…，$\sqrt[n]{\ }$表出诸x_i的一般解。在欧拉时代以前，人们已经知道对带有文字系数的二次方程、三次方程及四次方程而言（相当于$n=2$，3，4的情形），确实能利用系数的四则运算与开方根运算作成的公式表示出各个根。

既然代数方程中的多项式是由x本身和它的各次乘方x^2，…，x^n构成的，求根是个逆过程，所以理所当然地在计算根的过程中将会包含各次开方运算，即$\sqrt{\ }$，…，$\sqrt[n]{\ }$等运算，而且正是这些开方根$\sqrt{\ }$，…，$\sqrt[n]{\ }$的多值性，将能用以区别各个根。

试以α表$\sqrt[n]{1}$的一个复数根（幅角最小的复数），则共有n个

复数根α，α^2，…，α^{n-1}，$\alpha^n=1$。特别$\sqrt[2]{1}$的两个平方根为-1，$+1$，$\sqrt[3]{1}$的三个立方根为ω，ω^2，$\omega^3=1$，其中

$$\omega=-\frac{1}{2}+\frac{\sqrt{3}}{2}i,\ \omega^2=-\frac{1}{2}-\frac{\sqrt{3}}{2}i,\quad (i=\sqrt{-1}).$$

因此n次方程$f(x)=0$的根式解问题，就是如何利用a_0，a_1，…，a_n诸系数以及各次单位方根±1，ω，ω^2，…，α，α^2，…，α^{n-1}的有理运算及各次开方运算$\sqrt{\ }$，$\sqrt[3]{\ }$，…，$\sqrt[n]{\ }$来表示出各个x_k根的问题。

拉格朗日精心分析了二次、三次、四次方程的根式解结构之后，发现方程的预解式概念，并且还进一步看出预解式和诸根排列置换下的形式不变性有关。因此他最后得出结论说：高次代数方程根式解的可能性问题最终还是归结到诸根的排列置换性质的问题。

什么叫做拉格朗日预解式？我们不妨取二次方程与三次方程为例来说明问题。

大家知道二次方程$x^2+bx+c=0$的求根公式是

$$x_1=\frac{1}{2}\left(-b+\sqrt{b^2-4c}\right),\quad x_2=\frac{1}{2}\left(-b-\sqrt{b^2-4c}\right).$$

在这个公式中果然只用到了系数间的有理运算与开方$\sqrt{\ }$运算，而且借助于$\sqrt{1}=\pm1$的双值性区分了两个根x_1与x_2。那么$\sqrt{b^2-4c}$是怎样形成的呢？它与诸根有何关系？很显然，这个运算式可表成$\sqrt{b^2-4c}=x_1-x_2$。简记为$\Phi(x_1,x_2)=x_1-x_2$，则Φ即称为二次方程的预解式。由韦达关系已知

$$\begin{cases}x_1+x_2=-b\\x_1\cdot x_2=c\end{cases}$$

所以只要能将Φ用b，c表出，则由$x_1+x_2=-b$立即可得二根

$$x_1=\frac{1}{2}(-b+\Phi),x_2=\frac{1}{2}(-b-\Phi).$$

显而易见$\Phi^2=(x_1-x_2)^2=x_1^2+x_2^2-2x_1x_2=(x_1+x_2)^2-$

$-4x_1x_2 = (-b)^2-4c=b^2-4c$。故 Φ满足二项方程

$$\Phi^2-(b^2-4c)=0。$$

这就是二次方程的预解方程。由此得

$$\Phi=\sqrt{b^2-4c}\ 。$$

由上可见，只要求得预解方程，则原方程的根式解即随手可得。下面再来考察三次方程的根式解问题。

试考虑既约三次方程$x^3+qx+p=0$。它的三个根与系数的关系是

$$\begin{cases} x_1+x_2+x_3=0 \\ x_1x_2+x_2x_3+x_3x_1=q \\ x_1x_2x_3=-p\ 。\end{cases}$$

根据16世纪意大利数学家塔塔里亚(Tartaglia)与卡丹(Cardan)发现的解法，首先应作变换$x=y-\frac{1}{3}q/y$。代入原方程，经过简化则得

$$y^6+py^3-\left(\frac{q}{3}\right)^3=0。\tag{1}$$

这可简记为

$$\Phi^2+p\Phi-\frac{1}{27}q^3=0,\tag{2}$$

其中$\Phi=y^3$。由于这是一个二次方程，立即可解出Φ，最后再经过开立方可求得y的六个值：

$$y=\sqrt[3]{\Phi_1}\ ,\ \omega\sqrt[3]{\Phi_1}\ ,\ \omega^2\sqrt[3]{\Phi_1}\ ;$$

$$\sqrt[3]{\Phi_2}\ ,\ \omega\sqrt[3]{\Phi_2}\ ,\ \omega^2\sqrt[3]{\Phi_2}\ .$$

这里Φ_1与Φ_2是关于Φ二次方程的两个根，即

$$\Phi_i=-\frac{p}{2}\pm\sqrt{\left(\frac{p}{2}\right)^2+\left(\frac{q}{3}\right)^3},\quad (i=1,2)。$$

最后，经过恰当搭配可把原三次方程的三个根明确地表示如下：

$$x_1=\sqrt[3]{-\left(\frac{p}{2}\right)+\sqrt{\left(\frac{p}{2}\right)^2+\left(\frac{q}{3}\right)^3}}$$
$$+\sqrt[3]{-\left(\frac{p}{2}\right)-\sqrt{\left(\frac{p}{2}\right)^2+\left(\frac{q}{3}\right)^3}},$$
$$x_2=\omega\cdot\sqrt[3]{-\left(\frac{p}{2}\right)+\sqrt{\left(\frac{p}{2}\right)^2+\left(\frac{q}{3}\right)^3}}$$
$$+\omega^2\cdot\sqrt[3]{-\left(\frac{p}{2}\right)-\sqrt{\left(\frac{p}{2}\right)^2+\left(\frac{q}{3}\right)^3}},$$
$$x_3=\omega^2\cdot\sqrt[3]{-\left(\frac{p}{2}\right)+\sqrt{\left(\frac{p}{2}\right)^2+\left(\frac{q}{3}\right)^3}}$$
$$+\omega\cdot\sqrt[3]{-\left(\frac{p}{2}\right)-\sqrt{\left(\frac{p}{2}\right)^2+\left(\frac{q}{3}\right)^3}}.$$

这些公式果然也只用到了系数间的有理运算与平方根、立方根运算，且诸根的区别是借助于$\sqrt{1}$的双值性及$\sqrt[3]{1}$的三值性表示出来的。

我们已经看出，导出上述公式的重要关键是由于得到了关于Φ的二次方程，这个二次方程就是关于三次方程根式解法过程中的拉格朗日**预解方程**。这个预解方程式的得来是由于作了变数代换$x=y-\frac{1}{3}(q/y)$。

对这样一个变数代换而言，拉格朗日不是着眼于x如何依赖于y，而是考虑y如何依赖于x的问题。反过来分析问题，这便是他的高明之处。

拉格朗日观察到各个y值可以表成

$$y=\frac{1}{3}(x_1+\omega x_2+\omega^2 x_3) \tag{3}$$

这里三个字母x_1，x_2，x_3共有$3!=6$个排列，故对应各个排列可得y的6个值，所以y理应满足一个6次方程。

再者，对诸x_k的6种排列而言，

$$(x_1+\omega x_2+\omega^2 x_3)^3 \tag{4}$$

实际上只给出两个不同形式的值，这也说明为什么$y^3=\frac{1}{27}(x_1+\omega x_2+\omega^2 x_3)^3=\Phi$恰好满足一个两次方程——预解方程，其中$\Phi$就是拉氏预解式。

上述分析表明预解方程的次数与字母（根）x_1、x_2，x_3的排列所导致的不变形式的个数有关。为详细地说明这一情况。可把6个排列所对应的一切置换简记如下（可用$\mathfrak{S}_3$表示这个置换类）：

$$\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}\ \begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}\ \begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}$$

$$\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}\ \begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}\ \begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}$$

例如上列第三个置换即表示$x_1\to x_2$，$x_2\to x_1$，$x_3\to x_3$。最后一个置换即表示$x_1\to x_3$，$x_2\to x_2$，$x_3\to x_1$。

显然对上述6个置换而言，y的不同形式值也有6个。但在$(x_1+\omega x_2+\omega^2 x_3)^3$上施行6个置换时，只有两个不同的形式值，即

$$(x_1+\omega x_2+\omega^2 x_3)^3\text{与}(x_2+\omega x_1+\omega^2 x_3)^3。$$

事实上前者对下列3个置换是不变的：

$$\mathfrak{S}_{31}=\left\{\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix},\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix},\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}\right\}.$$

后者对6个中的另外3个置换是不变的。

所以在拉格朗日的分析研究中，已经发现了函数形式在一类置换下的形式不变性如何同预解方程的结构特征相关联。

循着上述思路，拉格朗日考察了更一般的情形，即假设x_1，x_2，…，x_n是方程

$$x^n+a_1x^{n-1}+\cdots+a_n=0 \tag{5}$$

的n个根，经过分析研究，他曾在1770年证明了两个重要命题。

熟知n个字母x_1，x_2，…，x_n的全排列共有$n!$个，故相应的全部置换构成的置换类$\mathfrak{S}_{n!}$中共有$n!$个不同置换。当然此类中的

一部分置换可作成置换子类。凡由方程系数a_1，…，a_n及字母x_1，…，x_n间的有理运算构成的表达式即称为有理式或有理函数。于是拉格朗日的两个重要命题可陈述如下：

（一）设$\Phi(x_1,x_2,\cdots,x_n)$和$\Psi(x_1, x_2, \cdots,x_n)$是两个有理函数（有理式）。如果使得$\Psi(x_1, x_2, \cdots, x_n)$不改变形式的某类置换$\mathfrak{S}$，也对$\Phi(x_1, x_2, \cdots, x_n)$不变形，则$\Phi(x_1, x_2, \cdots, x_n)$必可用$\Psi(x_1, x_2, \cdots, x_n)$及(5)的系数$a_k$ $(k=1, \cdots, n)$的某种有理形式表示出来。

（二）对使$\Psi(x_1, x_2,\cdots,x_n)$不变形的置换类$\mathfrak{S}$而言，如果$\Phi(x_1, x_2, \cdots, x_n)$取$r$个不同形式值，则$\Psi$必为某个$r$次方程之根，该方程之系数可用$\Psi$及（5）之系数的有理形式表示出来。

根据上述二命题，拉格朗日曾设想了一个理论上（原则性）的求解方程步骤：

例一 考虑二次方程 $x^2+bx+c=0$ 的求解问题。设x_1，x_2为方程的两个根，则有置换类：

$$\mathfrak{S}_2: \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

按根与系数的关系知

$$x_1+x_2=-b, \quad x_1\cdot x_2=c.$$

这些关系式都对置换类$\mathfrak{S}_2$不变形，所以它们能用原二次方程的系数表出，是理所当然的（这里不妨假定$\Psi\equiv1$）。

今若选取$\Psi(x_1, x_2)=x_1+x_2$，而考虑$\Phi(x_1, x_2)=x_1-x_2$，则易见Φ对使Ψ不变形的置换类$\mathfrak{S}_2$而言将取两个同形式值：x_1-x_2与x_2-x_1。因此由上述基本命题，Φ必为某二次方程之根，事实果然如此，即

$$(x_1-x_2)^2=(x_1+x_2)^2-4x_1x_2=b^2-4c,$$

也即 $$\Phi^2-(b^2-4c)=0.$$

由此，引入这个预解方程的根式解$\sqrt{b^2-4c}$，利用$x_1-x_2=\sqrt{b^2-4c}$ 与$x_1+x_2=-b$便得到

$$x_1=\frac{1}{2}(-b+\sqrt{b^2-4c}).$$

例二 试考虑三次方程$x^3+qx+p=0$，并令x_1，x_2，x_3表示该方程的三个根，令取

$$\Phi\equiv\Phi(x_1,x_2,x_3)=\frac{1}{27}(x_1+\omega x_2+\omega^2 x_3)^3,$$

由于这个有理式Φ对$\mathfrak{S}_3$中的六个置换而言，它只取两个不同值A和B，因此如取

$$\Psi(x_1, x_2, x_3)=x_1+x_2+x_3=0$$

则由基本命题（二），可知作为Φ的二值A，B必为某二次方程（预解方程）之根。于是再从下列诸式

$$\begin{cases}\frac{1}{3}(x_1+x_2+x_3)=0\\ \frac{1}{3}(x_1+\omega x_2+\omega^2 x_3)=\sqrt[3]{A}\\ \frac{1}{3}(x_1+\omega^2 x_2+\omega x_3)=\sqrt[3]{B}.\end{cases}$$

即可直接解出x_1，x_2，x_3。

拉格朗日还利用类似的思想方法处理了四次方程的根式解问题。可是用上述想法处理五次方程时却未获成功（欲知其详，可参阅Kline《古今数学思想》一书的第25章）。

但是拉格朗日的基本命题和方法还是有用的。鲁飞尼就是利用了拉氏方法证明了一个重要定理：当方程次数$n>4$时，不存在n个元的有理函数$\Phi(x_1, \cdots, x_n)$使得在n个元素的置换下能取3个或4个相异值（这就意味着不可能存在低于5次的预解方程）。由此，他便进一步大胆论证：任何高于4次的一般代数方程没有根式解。

可是鲁飞尼的证明是难以令人信服的，因为他在证明中用到了一个未加证明的辅助命题，后称**阿贝尔定理**。

阿贝尔定理是说，如果一个代数方程能用根式求解，则在根

式解法公式里出现的各个根式（如$\sqrt{\ }$，$\sqrt[n]{\ }$，等），一定都可以表示成方程诸根及某些单位根的有理式。

例如，二次方程$ax^2+bx+c=0$（$a\neq 0$）根式解法公式中出现的$\sqrt{b^2-4ac}$可表为 $a(x_1-x_2)$。又既约三次方程$x^3+qx+p=0$ 根式解法公式中出现的某个$\sqrt[3]{\ }$及$\sqrt{\ }$可分别表示成

$$\sqrt{-\left(\frac{p}{2}\right)+\sqrt{\left(\frac{p}{2}\right)^2+\left(\frac{q}{3}\right)^3}}=\frac{1}{3}(x_1+\omega x_2+\omega^2 x_3),$$

$$\sqrt{\left(\frac{p}{2}\right)^2+\left(\frac{q}{3}\right)^3}=\frac{1}{27}(x_1+\omega x_2+\omega^2 x_3)^3-\frac{1}{2}x_1x_2x_3.$$

只要明确根式代数解的全部要求，即不难从直观上理解上述阿贝尔定理的合理性。但其仔细证明还是需要不少笔墨的（自然韦达关系和拉氏重要命题等都是证法中需要借助的工具）。

阿贝尔正是应用上述定理才精确地证明了高于四次的一般代数方程不能有一般根式解的著名结论的。

§3 伽罗瓦的思想方法

在人类的数学发展史上，至少有三大发明是值得大书特书的，一是解析几何的发明，在数学概念思维领域实现了**形数关系的沟通**；二是微积分的发明，使人脑形式思维**进入无限小分析领域**；三是群论的发明，使代数研究进入新时代，即从局部性研究转向**系统结构的整体性分析研究阶段**。

伽罗瓦是群论思想方法的重要创始人之一，且是首先利用群论方法完成代数方程可解性理论的数学家。现在就专来谈谈伽罗瓦的基本思想方法。

对具有根式可解性的代数方程的特征问题，原来阿贝尔已在进行深入研究。不幸的是1829年死神夺去了他的生命，因而使即将完成的光辉事业功亏一篑。伽罗瓦的遭遇比阿贝尔更为不幸。在短促的一生中既未能考进他所向往的大学，也没有机会在他生

前发表他的重大创造性成果。他的遗稿是在他临死前一夜匆忙写成后委托他的朋友薛伐里叶（Chevalier）保存下来的。他因参加无意义的决斗死于1832年5月31日。直到1846年，法国数学家刘维耳才把伽罗瓦的科研成果整理后首次发表于刘维耳主编的数学杂志（Journal de Mathe′matiques）上。见该刊第11卷 381－444页。自此以后，伽罗瓦的重大贡献才逐渐为人们所习知。

从上一节的讨论已经看到，代数方程的根式解法公式是具有“层次式结构”的。例如二次方程求根公式中包含一层平方根，三次方程求根公式中出现两层根号，里边一层是平方根式，外边一层是立方根式。一般说来，一个高次代数方程如果存在根式解公式，则公式中必将包含由开方根运算构成的一些层次。

为了彻底解决方程根式解难题，伽罗瓦的基本思想方法要点可归结为如下三点：

（1）把**层次式结构**的形式同**域的不断扩张**概念联系起来。

（2）把每一层次的**对应域**的形成要素归结为**预解式**和**预解方程**的寻求。

（3）把预解式的寻求归结为**置换群**的各阶**子群**的结构分析。

特别重要的是上述第（3）点。

我们知道，对一般代数方程来说，关于拉氏预解式的构成并不存在明确的方法或法则，而对于较特殊的方程要构造预解式也需要很大的技巧（我们从三次方程根式解分析中已经看到了这一点）。因此，伽罗瓦的主要思想就是要设法绕过拉氏预解式。

伽罗瓦的重要功绩就在于他对置换群引进了一些重要概念（如“正规子群”、“单群”、“复群”以及群与群之间的“同构”等概念），并且还证明了一些基本定理。这样他终于成功地克服了拉氏预解式带来的困难，顺利地建立了方程的**可解性理论**。

我们知道，置换可以相继进行。例如

$$\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}\cdot\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}=\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}\cdot\begin{pmatrix}2&1&3\\1&3&2\end{pmatrix}=\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}.$$

不妨把“点乘”看作是置换间的乘法。一般来说，这种乘法未必满足交换律，例如

$$\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}\cdot\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}=\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}\cdot\begin{pmatrix}3&1&2\\3&2&1\end{pmatrix}=\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}.$$

故 $\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}\cdot\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}\neq\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}\cdot\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}.$

但不难验明三个置换相乘时，恒满足结合律（即$(\theta_1\cdot\theta_2)\cdot\theta_3=\theta_1\cdot(\theta_2\cdot\theta_3)$）。而且不动置换可当作乘法单位元，从而还有乘法逆元存在。例如就$\mathfrak{S}_6$中的六个置换而言，

$$\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}\cdot\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}=\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix},\quad \begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}^{-1}=\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix};$$

$$\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}\cdot\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}=\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix},\quad \begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}^{-1}=\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix};$$

$$\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}\cdot\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}=\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix},\quad \begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}^{-1}=\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix};$$

$$\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}\cdot\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}=\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix},\quad \begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}^{-1}=\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix};$$

等等。易看出乘法左逆和右逆是同一的。比如

$$\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}^{-1}\cdot\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}=\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}\cdot\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}^{-1}=\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}.$$

所以对“点乘·”而言$\mathfrak{S}_6$构成一个群，这种群叫作置换群。

一般说来，n个根作成的全排列给出置换群$\mathfrak{S}_{n!}$。这种群还可能包含着子群。例如§2中说到的$\mathfrak{S}_3$就是$\mathfrak{S}_6$的一个子群。又只包含一个不动置换的单元集$\mathfrak{S}_1$也是$\mathfrak{S}_6$与$\mathfrak{S}_3$的子群，这是任何置换群都有的最小子群。

如所周知，所谓群的阶就是群中的元素（例如置换群中的置换）的个数。故$\mathfrak{S}_6$，$\mathfrak{S}_3$，$\mathfrak{S}_1$，的阶数分别为6，3，1。拉格朗日的一个简单定理是说，群的阶一定是子群阶的倍数。这种倍数称为该子群的**指标**。例如在$\mathfrak{S}_6$中的子群$\mathfrak{S}_3$具有指标6/3＝2，可记作

$[\mathfrak{S}_6/\mathfrak{S}_3]=2$. 同理$[\mathfrak{S}_3/\mathfrak{S}_1]=3$。

设G是一个群，H为G的子群，则H的指标即可记作$[G/H]$。如果对G中的每一元素g均有$g\cdot H=H\cdot g$，则称H为G的**正规子群**或**不变子群**。这里$g\cdot H$表示g与H中诸元素h的乘积$g\cdot h$构成的集合。同理，$H\cdot g$表示集合$\{h\cdot g\}$。故$g\cdot H=H\cdot g$即表示$\{g\cdot h\}$与$\{h\cdot g\}$（$h\in H$）是同一个集合（利用H还可将G划分为$[G/H]$个"陪集"）。

不难验明，$\mathfrak{S}_3$是$\mathfrak{S}_6$的正规子群，而$\mathfrak{S}_1$是任何置换群的正规子群。后面将会看到，正规子群的概念在伽罗瓦理论中起着极为重要的作用。

现在我们来按照伽罗瓦的想法重新考察一次、二次、三次方程的一般代数解问题。从这里将可看出他的普遍思想方法是如何以拉氏重要命题（二）和阿贝尔定理作为基础发展而成的。

一次方程$ax+b=0$只有一个根x_1，置换只有一个，即$x_1\to x_1$，故相应的置换群为单元群$\mathfrak{S}_1=\left\{\begin{pmatrix}1\\1\end{pmatrix}\right\}$。今将系数$a$，$b$加入到通常的有理数域$R$中进行有限次有理运算的全部结果构的域$[a,b;R]$叫作伽罗瓦**有理域**。用这个有理域中的元素即可解出方程.这样，就说该方程在该有理域上是可解的。事实上，$x_1=b/a$（$a\neq 0$），b/a确实属于该有理域。

二次方程$x^2+bx+c=0$有两个根x_1，x_2，相应的置换群为$\mathfrak{S}_2=\left\{\begin{pmatrix}1&2\\1&2\end{pmatrix},\begin{pmatrix}1&2\\2&1\end{pmatrix}\right\}$，$\mathfrak{S}_1=\left\{\begin{matrix}1&2\\1&2\end{matrix}\right\}$，子群$\mathfrak{S}_1$的指标为$[\mathfrak{S}_2/\mathfrak{S}_1]=2$.但是光利用$b$，$c$与$R$作成的有理域$[b,c;R]$中的元素还不足以把$x_1$,$x_2$解出来，原因是

$$x_1+x_2=-b,\qquad x_1\cdot x_2=c,$$

这里两根的代数有理式x_1+x_2与$x_1\cdot x_2$对$\mathfrak{S}_2$中置换都不变形.换言之,在$\mathfrak{S}_2$的全体置换范围内是不能把x_1和x_2区分开来的。为了解出各个x_k（$k=1,2$），就必须进行区分。于是考察只在较小的

置换范围内不变形的代数有理式。不妨把置换范围取成子群$\mathfrak{S}_1$。对这个子群中的置换而言，不改变形式的有理式中的最简单者莫过于x_1-x_2。采用伽罗瓦的说法，式x_1-x_2所相应的不变子群即$\mathfrak{S}_1$，而对$\mathfrak{S}_1$之外的置换（如$\begin{pmatrix}1 & 2\\ 2 & 1\end{pmatrix}$）则改变形式，即由$x_1-x_2$改变为$x_2-x_1$。因此，根据拉格朗日的重要命题（二）可知$\Phi\equiv x_1-x_2$必是某二次方程之根。事实上，预解式$\Phi$所满足的方程次数2是指标$[\mathfrak{S}_2/\mathfrak{S}_1]=2$决定的（因为$\mathfrak{S}_2$按$\mathfrak{S}_1$划分为“陪集”时恰好有两个陪集，对不同陪集中的置换，Φ将取不同的形式。故Φ取不同值的个数2即由$[\mathfrak{S}_2/\mathfrak{S}_1]$所给出）。

果然，$\Phi^2=(x_1-x_2)^2=(x_1+x_2)^2-4x_1x_2=b^2-4c$。由此$\Phi=\sqrt{b^2-4c}$。因此，为明确解出$x_1$与$x_2$，还必须将这个代数量加进到原来的有理域$[b,c,\sqrt{b^2-4c};R]$。利用这个扩张域中的元素果然足以把$x_1$，$x_2$表示出来了，即得熟知的二次方程求根公式。

上述分析表明，为了获得求根公式，应将公式中的根式层次结构的形成同有理域的扩张概念联系起来。域是通过加进新的代数量而形成扩张的，而新的代数量（表现为根式的量）正好是求根公式结构方式中的一个层次。如用建造房屋作比喻，则伽罗瓦域好比是提供造房器材的仓库。如果只建造最简易的民用土房，则仓库里只要有砖瓦木料就够了（例如，解一次方程时，只要有理域$[a, b; R]$就够了）。但要建造一层楼时，原有的简易仓库就不够用了，至少需要加进“预制板”或“水泥板”等器材才行。因此仓库的内存需要扩大。这就好比是$[b, c; R]$必须扩张成伽罗瓦有理域$[b,c,\sqrt{b^2-4c};R]$才足以构造二次方程求根公式一样。

由上述已知代数量$\sqrt{b^2-4c}$是通过预解方程得到的，所以首先要找出预解式$\Phi=x_1-x_2$。伽罗瓦的突出贡献，就是直接依靠置换群的子群结构分析来预见预解式的形成方式及其应满足的预解方程的类型。例如，通用$\mathfrak{S}_2$的正规子群$\mathfrak{S}_1$及$[\mathfrak{S}_2/\mathfrak{S}_1]=2$即

可自然地找出$\Phi=x_1-x_2$及其应满足的二次方程（实则为二项方程）。

一次、二次方程毕竟太简单了，似乎不值得"杀鸡用牛刀"。下面再来分析三次方程求根公式是如何产生的。

令x_1，x_2，x_3为三次方程$x^3+qx+p=0$的三个根。则等价于该三次方程的韦达关系式中的诸对称函数具有不变置换群$\mathfrak{S}_6$。已知$\mathfrak{S}_3$是$\mathfrak{S}_6$的正规子群，而$\mathfrak{S}_1$是$\mathfrak{S}_3$的正规子群。相应的指标依次是：

$$[\mathfrak{S}_6/\mathfrak{S}_3]=2,\ [\mathfrak{S}_3/\mathfrak{S}_1]=3.$$

按伽罗瓦的方法，不必象拉格朗日那样去硬凑一个有理函数Φ使其在$\mathfrak{S}_6$的一切置换下恰好取两个不同值，而是直接去寻找一个在$\mathfrak{S}_3$置换下不变形的有理式（这样作，工作量少得多了）。由观察或试验（或经过系统性的试算）可知

$$\Phi(x_1,\ x_2,\ x_3)=\frac{1}{27}(x_1+\omega x_2+\omega^2x_3)^3,$$

正好具有最大的不变置换群$\mathfrak{S}_3$(注意系数$\frac{1}{27}$的附上是无关紧要的，只是为了写预解方程时形式可以简明一些)。再者，易见

$$\Phi_1(x_1,x_2,x_3)=\frac{1}{3}(x_1+\omega x_2+\omega^2x_3)$$

具有不变置换群$\mathfrak{S}_1=\left\{\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}\right\}$.

既然$\mathfrak{S}_3$与$\mathfrak{S}_1$的指标分别为2与3，故可知Φ与Φ_1应分别满足一个二次方程与三次方程但后者应通过Φ及方程系数的有理形式表示出来。事实上，关于Φ与Φ_1的预解方程分别为

$$\Phi^2+p\Phi-\frac{1}{27}q^3=0,\ \Phi_1^3-\Phi=0.$$

通过求Φ，知有理域$[p,\ q;\ R]$应扩张为

$$\left[p,q,\sqrt{\left(\frac{p}{2}\right)^2+\left(\frac{q}{3}\right)^3};\ R\right],$$

再通过求Φ_1，可知上述伽罗瓦有理域还须进一步扩张为

$$\left[p, q, \sqrt{\left(\frac{p}{2}\right)^2+\left(\frac{q}{3}\right)^3}, \sqrt{-\left(\frac{p}{2}\right)\pm\sqrt{\left(\frac{p}{2}\right)+\left(\frac{q}{3}\right)^3}}; R\right].$$

这样一来，在这个域内果然可将原方程的三个根都明确地表示出来了（参阅§2有关内容）。

这里补充解释一下，为什么不变子群的指标恰好等于预解方程的次数呢？事实上，这是因为置换群可按子群划分成陪集的并集，陪集的个数等于该子群的指标。如果该子群正好是某有理函数Φ的不变置换群，则Φ在各个陪集的置换下将取不同形式值，不同形式值的个数与陪集个数相同，因而与该子群的指标相同。故根据拉氏重要命题（二），可知指标应等于Φ所满足的方程的次数。

§4 方程式可解性理论简介

从上一节关于三次方程的讨论中我们已看到，置换群与伽罗瓦有理域是相互对应的，例如，群$\mathfrak{S}_6$对应于$[p, q; R]$，子群$\mathfrak{S}_3$对应于有理域

$$\left[p, q, \sqrt{\left(\frac{p}{2}\right)^2+\left(\frac{q}{3}\right)^3}; R\right]$$

等。特别，$[p, q; R]$称为方程系数的有理域，其对应的置换群$\mathfrak{S}_6$即称为方程的**伽罗瓦群**。

今将伽罗瓦所设计的一般方法步骤概述如下：(1) 对于给定的一个一般的或特殊的含有未知根$x_1, \cdots, x_n$的代数方程，首先找出同该方程系数有理域F相应的伽罗瓦群G。(2) 设法找一个G的最大子群H，如果阶数相同的最大子群不只是一个，则可任取其一。(3) 找到最大子群H之后，则可通过一系列有理代数运算的手续，设法找到一个函数$\Phi \equiv \Phi(x_1, \cdots, x_n)$，其系数属于$F$，且对置换群$H$中的一切置换恒保持形式不变，而对$G$中不在$H$内的其它置换则改变其形式值。换言之，要使$H$恰好成为$\Phi$

的最大的不变置换群。事实上，很可能存在无穷多个如上所说的函数，但只须选定其中Φ即可。（4）如可能，应设法在域F内构造出一个以Φ作为其一根的代数方程式，其次数应等于指标$[G/H]$。作出的方程就称为预解方程。（5）要从预解方程中解出Φ，从而将获得一种新的代数量（用根式表示的量），然后再将这个表现Φ的量加入到域F中去，得到扩张了的有理域$F'\supset F$。对应于这个扩张域F'的伽罗瓦群就是H。

接下去再从H出发，重复上述过程。最后便可得到一个最大扩张的有理域，它所对应的伽罗瓦群正好是由不动置换作成的单元群$E=\left\{\begin{pmatrix}1&2&\cdots&n\\1&2&\cdots&n\end{pmatrix}\right\}$。于是原方程的$n$个根便可由最终得到的最大扩张域中的元素表示出来。这样便找到了代数方程的根式解。

但是要注意，并非对任何高次方程上述过程总是能行得通的。事实上，上述过程中的步骤（4）不一定总能实现，因为预解方程未必总能找到。因此，还必须对方程的根式可解性建立一个判别准则（这种准则正是阿贝尔生前曾努力寻求过的一个未知目标）。

伽罗瓦在精细地研究了拉格朗日、高斯、阿贝尔、柯西等人著作的基础上，终于发现了一个重要准则，这就是如下所说的伽罗瓦基本定理：

给定一个代数方程，假设G是该方程的伽罗瓦群（即关于方程诸根的某个置换群），它的一系列最大**正规子群**为H_k（$k=1,\cdots,s$）：

$$G\supset H_1\supset\cdots\supset H_s,\quad (H_s\text{为单元群})。$$

则原方程可用根式求解的**必要与充分条件**是，下列诸指标

$$[G/H_1],[H_1/H_2],\cdots,[H_{s-1}/H_s]$$

都是素数。

事实上，伽罗瓦曾示明当且仅当H_1为正规子群而$[G/H_1]$为素数p时，则相应的预解方程才可写成或约化成形如$x^p=A$的二

项方程（A属于G相应的有理域），从而根据高斯定理，x可用根式解出。因为原方程可否用根式求解的问题，恒可归结为是否存在一系列形如$x^p=A$的二项方程（即预解方程）的问题，所以伽罗瓦基本定理的合理性是容易理解的。

由上可知，伽罗瓦所以能给出上述基本定理的重要关键，是由于他发现了正规子群这一重要概念，否则代数方程根式解的充要条件或判别准则就无法表述出来。

有了根式可解性判别准则之后，再回过头来检验鲁飞尼-阿贝尔关于高于四次的一般方程不存在根式解的结论就非常简单了。

考虑一般n次代数方程。假设$n\geqslant5$。则相应的伽罗瓦群即置换群$\mathfrak{S}_{n!}$(又称n次完全对称群)，其阶为$n!$。这个群的最大正规子群正好是轮回子群$\mathfrak{S}_{n!/2}$其阶为$n!/2$。而这个轮回群的最大正规子群为单元群$\mathfrak{S}_1$。因此最大正规子群的组成列为

$$\mathfrak{S}_{n!}\supset\mathfrak{S}_{n!/2}\supset\mathfrak{S}_1.$$

相应的指标数列为2，$n!/2$。

注意在$n>4$时，$n!/2$恒为一合数而不是素数，故根据伽罗瓦基本定理立即推知高于四次的一般代数方程不可能有根式解。

另一方面，对于一般的二次方程及三次方程而言，它们分别具有组成列：

$$\mathfrak{S}_2\supset\mathfrak{S}_1=\left\{\begin{pmatrix}1&2\\1&2\end{pmatrix}\right\},$$

$$\mathfrak{S}_6\supset\mathfrak{S}_3\supset\mathfrak{S}_1=\left\{\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}\right\}.$$

相应的指标数列分别为2和2，3。这些都是素数，所以它们都具有根式解。

对于一般四次方程，可以验证它具有组成列：

$$\mathfrak{S}_{24}\supset\mathfrak{S}_{12}\supset\mathfrak{S}_4\supset\mathfrak{S}_2\supset\mathfrak{S}_1=\left\{\begin{pmatrix}1&2&3&4\\1&2&3&4\end{pmatrix}\right\}.$$

相应的指标数2，3，2，2也都是素数。故按可解性判别准则立知

四次方程也有根式解。事实上，早在16世纪40年代意大利的费拉里（Ferrari）就已经发现四次方程的根式解了。

一般说来，如果一个代数方程具有根式解，则由相应的指标数列还可预断根式解过程中需要求解相应次数的二项方程。例如，就四次方程而言，指标数列2，3，2，2即表明根式解过程中需要求解四个二项方程，其中一个是三次的二项方程，三个是二次的二项方程。

关于方程可解性理论（又名伽罗瓦理论）的简略介绍即到此为止。最后，对伽罗瓦思想方法发展过程中的重要关键之处似乎还值得重新分析一下。从§2～§4三节的内容看来，不难发现，伽罗瓦的研究思路是在拉格朗日、阿贝尔、高斯等先驱者工作成果的启发下形成的。置换的概念、数域的概念、根式求解必然归结到二项方程等想法，早在他的先驱者那里就已经形成了。假如阿贝尔不是不幸而夭折，则阿贝尔也很可能会通过同样的道路去建立可解性理论。“英雄所见略同”，实际上他们的思想方法已经非常相似。

伽罗瓦从拉格朗日那里继承了问题转化的思想，即把预解式的构成同置换群联系起来的思想，并且发展了这一思想。最后，伽罗瓦非常彻底地把全部问题都转化为或者归结为置换群及其子群结构分析的问题。这就形成了一个“突破”。

高斯早就预见到代数方程的根式解问题终归为二项方程的求解问题，并且对此提供了重要成果。伽罗瓦仔细分析了具有根式解的二项方程作为“预解方程”时所相应的置换子群的特征。结果发现对应的置换子群必须是正规子群且指标应是素数才行（事实上只有这样的正规子群所对应的预解方程才可能是素数次的二项方程，从而能用根式求出预解式）。引出正规子群的概念并弄明白这种群的性质和作用，这是伽罗瓦工作中的第二个重大突破。

正是有了上述一些“突破”，才使得伽罗瓦终于作成了代数方

程的一般形式的“可解性理论”。继承和发展是科学研究发展的普遍规律，伽罗瓦的杰出贡献，显然也是符合这个普遍规律的产物。通常，当说到年轻的伽罗瓦发明群论的故事时，习惯地把他赞誉为头等天才，而并不强调他从杰出的先驱者那里努力继承了最宝贵的遗产，这给人们的印象自然是不够全面的。

第7讲 关于非标准数域与非康托型自然数模型的构造方法

§1 略论“无限”概念蕴含的矛盾

无限概念是包含着矛盾的。这是自古以来哲学界与数学界众所周知的事实。

翻开数学史就可知道，象无限小与无穷大（包括超穷集与超穷基数等概念）引入数学分析领域，都是颇费周折的。其原因是由于无限概念中蕴含着矛盾的原故。人们对待矛盾的各个侧面可以采取不同的观点，在数学领域中处理这些矛盾还有各种不同的方法和主张，这样最终便导致近代直觉主义者与形式主义者以及逻辑主义者之间的分歧与对立。事实上，从哲学上说来，关于无限概念的两种分歧观点早在芝诺（Zeno）、亚里斯多德（Aristotle）时代就已经萌芽了。

人们早就发现，对待无限或无限过程可以有两种截然不同的、在概念上互相排斥的理解方式，一是把“无限”看成为永远在延伸着的(即不断在创造着的永远完成不了的)变程或进程。例如不断延伸的自然数列$1,2,3,\cdots,n,n+1,\cdots$ 就具有这样的性质。二是把无限对象看成为可以自我完成的过程或无穷整体。例如把自然数全体理解为一个真无限集合$\{1,2,3,\cdots,n,\cdots\}$。显然，这两种理解方式是彼此排斥的，因为前者把无限理解为永远不能完成的进程，而后者则把无限理解为可以完成的过程。可见，无限能否完成或形成整体，这是问题的关键所在。

在数理哲学上，把进程式的无限称为**潜无限**或**消极无限**（又名假无限或恶无限），把过程式的无限或完成了的无限称之为**真**

无限或**实无限**（又名**绝对无限**）.

自古以来哲学家与数学家中就分成两大派别，即消极无限派与实无限派。举例言之，象亚里斯多德、高斯、克朗内克、布劳威耳（Brouwer）、庞卡莱、外耳（Weyl）等以及现代直觉主义者属于消极无限派。莱布尼茨、黑格尔（Hegel）、康托、罗素、希尔伯特以及现代柏拉图主义者都属于实无限派。也有一些人做法上采用实无限派的方法而在思想上并不承认任何实无限存在性观点的，例如现代非标准分析创始人鲁宾生(A.Robinson)就是一个突出的例子。按鲁宾生的奇怪说法，无论在现实经验里或理性思维里都不存在实无限，但数学家为了进行数学工作还只好把并非存在的东西当作似乎是存在着的那样。当然鲁宾生的"二元论"是不足取的。

但是鲁宾生的非标准分析却是值得肯定的。鲁宾生等人利用超幂概念成功地构造出有序的非标准实数域，使分析学领域中久已废黜的实无穷小概念重新纳入合法地位；使得新型的分析学容许无穷小直接参加运算而变成更为灵活的数学工具。这在某种意义下彻底实现了当初莱布尼茨发明微积分时的原始理想；因此理应看作是一项重要贡献。

大家知道，标准实数系满足阿基米得公理，所以标准实数域中并不存在任何无穷小。要想在实数域中添进各种无穷小（所谓非标准数），就必须把一些等价的无限过程类定义成无穷小。但这种无限过程显然不能在原来的标准实数轴上来界定（因为实数轴上的极限过程对应的极限值仍然是标准数），所以只有借助于二维空间的无限过程类才能给出本质上的新对象即非标准实数。又因为这些非标准实数——无限小都是由**完成了的等价过程类**来定义的，所以它们理所当然地可以称之为实无限小概念。下一节讨论非标准数域 $^*R$ 的构造时，即可理解这里所说到的一切。

数学发展史指出，微积分发明后曾经历过神秘主义的发展阶段。所谓神秘主义无非是指微积分学中利用与实数相矛盾的无穷

小量进行计算竟能求得正确的结果，而且象微商、积分这样一些最基本的概念也得借助于实数系中并不存在的无穷小量来表述。由此看来，非标准数域的建立，正好表现为标准实数与非标准元素（无穷小量及无穷大量）的对立统一。换言之，$^*R$乃是一个包含着非标准数与标准数域R作为其组成部分的矛盾统一体，它是解决无穷小存在性矛盾的一种高一层次的数学结构。

为了阐明无穷大概念的矛盾本质，只须分析自然数序列概念即可。如所共知，正是由于对自然数无限性的理解方式不同，才形成了近代各个数学流派观点与主张的分歧，所以关于自然数无限性的剖析有着特别重要的意义。

前面说到潜无限与实无限两种相互对立的概念时曾指出："承认不承认无限能否完成或形成整体，这是问题的关键所在。"其实在认识论上最根本的关键问题还在于承认不承认人脑的思维能够模写运动和反映"飞跃"？

现代的反映论者是承认人脑的思维运动能够反映客观存在的飞跃过程的。根据这样的观点，也就自然能够接受自然数序列作成无穷集合的概念了。事实上，为了从直观上掌握自然数序列的整体性概念，可以设想在坐标轴上一个动点从坐标1处向原点0移动，而当动点达到0时，也就通过了无穷点集（数集）

$$\left\{1,\ \frac{1}{2},\ \frac{1}{3},\ \cdots,\ \frac{1}{n},\ \cdots\right\}$$

中的一切点，又因为$\frac{1}{n}$与自然数n作成一一对应，所以一切自然数这个概念也就被确定下来（详细论述，请再参见第9讲有关段落）。

在上述思维形式中，实际上思维在模写运动：在动点滑到0的过程中，该动点逐次走过坐标点$\frac{1}{n}$（$n=1,2,3,\cdots$），心像也就跟踪前进。既然运动的实质是**联结与过渡**，所以该动点必然会达到0点，而在这一时刻也就在人们思想里立即呈现了（完成了）

一切$\frac{1}{n}$的概念。这是思维运动里的一个飞跃，它正好对应地反映了运动中的一个**阶段性飞跃**，即动点坐标从非零数值变为零的那个飞跃。

康托引进最小的超穷序数ω时即已十分明确：自然数序列的形成过程包含着两个阶段：一是延伸（进展），二是穷竭。只有经过延伸到穷竭才彻底扬弃了有限性，完成了真无限过程。仔细说来，上述过程的真无限性是由"飞跃"形成的。因此，自然数序列的过程结构理应表示成下列形式：

$$\{1,2,3,\ (\cdots)_0,\ n,\ \ (\cdots)_1\}$$

在这个表示法中，n代表任意的自然数，$(\cdots)_0$表示着自然数的不断有限延伸，即量变阶段；而$(\cdots)_1$表示着那扬弃了**有限性重复发生现象**的飞跃阶段，即质变阶段。

由上已知，自然数的真无限过程是由飞跃段$(\cdots)_1$来完成的，所以$(\cdots)_1$中已经蕴含了真无限性。可是其中的成员（自然数）按照康托的观点看来又都是有限的序号数，有限序号数增长到无限就否定了序数自身的有限性，可见$(\cdots)_1$中之有**无限多的有限序号数**这一概念本身就隐含着矛盾。按照恩格斯《反杜林论》里的话来说，无限性是矛盾，而且是满含矛盾。无限性只能由有限的量来构成，这已经是一种矛盾，可是事实上就是如此。

事实上，**所有的有限序数**这一极为平常的概念即可引出隐藏在其中的某种悖论，**"所有"**这个思想规定表明：任给$M>0$，总有序数$n>M$；所以序数数列必然增长至无限，即不可能囿于任何有限的界限内，这就否定了序数恒保持**有限性**的规定。另一方面，**有限**这个思想规定表明：序数由小增大，不允许增大至无限，也即规定序数数列不能增长至无限，因而不可能完成**所有**，故又矛盾于"所有"这规定。正因为恩格斯早就洞察到这种矛盾，故上述悖论不妨叫作恩格斯的无限悖论。

怎样去解决关于自然数无限性的上述矛盾呢？看来唯一的途径就是应该去正确解释飞跃段——（…）$_1$的内部结构。从哲学上看，这需要弄清楚潜无限性如何转化（过渡）成实无限性的问题。但如果只停留在抽象概念分析里，这仍然是得不到彻底解决的。在这一讲里，我们将利用非标准数域的构造方式，引出一种**非康托型自然数模型**，它将对飞跃段（…）$_1$的内部结构给出一种具体刻画，从而能清楚地表明潜无限向实无限过渡的飞跃形式。这从某种观点看，我们给出的模型也可认为是对恩格斯悖论的一种具体解释。

我们要构造的非康托自然数序列模型**N**将包含无穷多个比康托序数ω为小的无限大自然数，这批非标准自然数都是由延伸、穷竭原则产生的潜在无限大元素，它们正好组成了整个飞跃段（…）$_1$。利用非标准分析方法还可以证明上述潜在无限大元素的个数至少和标准实数一般多。具体细节将在后面的§3中给出。

§2　非标准数域的构造方法

1960年美国数理逻辑学家鲁宾生曾运用现代数理逻辑方法和新成果，成功地示明了无限小的存在性，从而把实数域R扩展成包含着数不清的无限小和无限大的非标准数域$^*R$。$^*R$中的元素和R中的标准实数一样，可以进行四则运算并遵守三分律。由此在$^*R$上重新展开分析学的讨论，就构成了非标准分析的内容（细节请参考鲁宾生于1966年出版的著名著作“Non-standard Analysis”（已有中文译本））。

就实质而言，非标准分析与标准分析是等价的，即前者所能证明的命题也能为后者所证明，反之亦然。虽然如此，在分析计算方式上前者往往比后者直捷了当些。这是因为在标准分析中采用的极限论方法实际是分两步走的方法，即先是逐步近似，然后再取极限的方法；非标准分析法由于使无限小与无限大等非标准实

数直接参与运算，故本质上是一步走的方法。如果精通这种方法，则对发现新命题将会提供方便。

但是鲁宾生原来所使用的数理逻辑工具十分繁重，这对于一般分析学家看来不易引起兴趣。美国的梵·奥士达(Van Osdol)于1972年发表的一篇文章（载Amer.Math.Monthly,79卷355-363）中，曾参照日本学者高桥（Takahashi）在1970年夏天所描述的思想方法，具体地构造了一个较特殊的非标准实数域模型，即所谓超幂 $^*R$，或叫高桥模型。为了使推理更加简明和直观化，下面我们将采用等价过程量的表述形式来介绍这种模型的构造方法。

（一）过程量概念．试考虑以0,1,2,…为段落的所谓整点阶梯函数：

$$x(t)=x_n,\quad (n\leqslant t<n+1)\quad n=0,1,2,\cdots$$

其中参变量 t 反映 $x(t)$ 的变化过程，实数序列 $\{x_n\}$ 表示在不同的变化时刻 $x(t)$ 的变化程度和趋势。记 $x(t)$ 作成的数集为

$$\widetilde{x}=\{x(t)\mid 0\leqslant t<+\infty\}$$

并把 $\widetilde{x}$ 叫做过程量，特别，当 $x(t)=x_n\to 0$（或 ∞）时，则相应的过程量 $\widetilde{x}$ 叫做“趋零”（或“趋无穷”）过程量。又如果 $x(t)=c$（常数），则称相应的 $\widetilde{x}$ 为常过程量，可简记为 $\widetilde{c}$。

关于过程量之间的加、减、乘运算可定义如下：

$$\widetilde{x}\pm\widetilde{y}=\{x(t)\pm y(t)\mid 0\leqslant t<+\infty\},$$

$$\widetilde{x}\cdot\widetilde{y}=\{x(t)\cdot y(t)\mid 0\leqslant t<+\infty\}.$$

为定义除法运算，只须给出 $\widetilde{x}$ 的倒数概念即可。如果 $x(t)$（$0\leqslant t<+\infty$）不恒为零，则可引入整点阶梯函数 $y(t)$（$0\leqslant t<+\infty$）：

$$y(t)=\begin{cases}0 & \text{当}x(t)=0\text{时},\\ 1/x(t) & \text{当}x(t)\neq 0\text{时}.\end{cases}$$

于是由 $y(t)$ 引出的过程量 $\widetilde{y}$ 即可规定为 $\widetilde{x}$ 的倒数，可记作 $\widetilde{y}=1/\widetilde{x}$。

我们的目的是要利用“过程量等价类”的概念来界定广义实

数（非标准实数与标准实数），因此需要确定过程量之间的等价性概念。过程量是否等价以及是否具有某种性质P等等均须有一种判别标准，这种标准必须能使得由它判定的性质或命题恒符合逻辑上的排中律要求。下面我们就来建立所需要的判别标准。

（二）超滤集概念。标准自然数集$\{n\}$可记为N，这里我们假定0是属于N的最小自然数，即$N=\{0,1,2,\cdots,n,\cdots\}$。于是$\{x_n\}$便是幂集$R^N$中的一成元。显然幂集$R^N$中的一切成元也就相应地确定了一切过程量，原因是有如下的一一对应关系：

$$\tilde{x}\longleftrightarrow x(t)\longleftrightarrow\{x_n\},\ \{x_n\}\in R^N.$$

记$x_n=f(n)$，$y_n=g(n)$，$z_n=h(n)$等，则关于过程量$\tilde{x}$，$\tilde{y}$，$\tilde{z}$的各种性质表述可以转化为对$f(n)$，$g(n)$，$h(n)$的性质表述。

滤集定义　假设N上的某一子集类$\mathfrak{S}$具有下列三条性质：1°空集$\varnothing\notin\mathfrak{S}$，2°设$S_1\in\mathfrak{S}$，$S_2\in\mathfrak{S}$，则$S_1\cap S_2\in\mathfrak{S}$。3°若$S\in\mathfrak{S}$，且$S\subseteq T\subseteq N$，则$T\in\mathfrak{S}$。那么，称$\mathfrak{S}$为$N$上的一个滤集(Filter)。

例。今考虑N上的子集类：

$$\mathscr{F}=\{S|S\subset N,\ N-S=\text{有限集}\},$$

则易验明$\mathscr{F}$就是N上的一个滤集（证明十分简单，此处从略），这里$N-S$表示差集。

为了能对$\{f(n)\}$，$\{g(n)\}$等数列的性质给出一意性的判别标准，我们还须对上例中的滤集$\mathscr{F}$尽可能地加以扩大使之成为一个符合下列定义的所谓超滤集：

超滤集定义　如果滤集$\mathscr{U}$具有这样的性质：对每一子集$S\subseteq N$，或者是$S\in\mathscr{U}$，或者是$(N-S)\in\mathscr{U}$，且两者必居其一，则称此$\mathscr{U}$为一超滤集。

存在定理　至少存在一个N上的超滤集以$\mathscr{F}$为子集。

证明　为证明上述定理，需要用到关于半序集（偏序集）的仑氏引理(Zorn's lemma)。试取

$$\mathfrak{C}=\{N\text{上的滤集}\mathscr{F}'|\mathscr{F}'\supseteq\mathscr{F}\},$$

即$\mathfrak{C}$为N上的所有包含$\mathscr{F}$的滤集所组成的类。显然$\mathfrak{C}$含有$\mathscr{F}$，故

$\mathfrak{C}$为非空。试注意$\mathfrak{C}$中诸滤集间的包含关系是一个偏序关系，并且$\mathfrak{C}$中任一上升的链在$\mathfrak{C}$中有上界，这个上界就是链中各项的并集。因此，由Zorn氏引理可知$\mathfrak{C}$中至少存在一个极大元$\mathscr{U}$。可以证明这个极大元$\mathscr{U}$正好就是N上的一个超滤集。

今用反证法证之。假设上述$\mathscr{U}$不是超滤集，亦即对此$\mathscr{U}$而言，条件"对每一$S\subseteq N$，或者有$S\in\mathscr{U}$，或者有$(N-S)\in\mathscr{U}$"并不成立。亦即存在一个非空子集$S\subseteq N$满足$S\notin\mathscr{U}$并且$(N-S)\notin\mathscr{U}$。从而可以推断必存在一个N的子集$T\in\mathscr{U}$使得$S\cap T=\varnothing$(空集)。为什么？要是不然的话，就是说：S与$\mathscr{U}$的每一元的交集**均非空**。但这样一来，即可证明

$\mathscr{U}'=\{X\subseteq N\mid X\supseteq S\cap T$，对某个$T\in\mathscr{U}\}\in\mathfrak{C}$，且$\mathscr{U}'\supset\mathscr{U}$。注意$S\cap T$非空，故$\mathscr{U}'$满足滤集条件1°，2°是显然的。这里只须验明条件3°：设$X_1\in\mathscr{U}'$，$X_2\in\mathscr{U}'$，则$X_1\cap X_2\in\mathscr{U}'$。事实上，由$\mathscr{U}'$定义知$X_1\supseteq S\cap T_1$，又$X_2\supseteq S\cap T_2$，其中$T_1\in\mathscr{U}$，$T_2\in\mathscr{U}$，故$T_1\cap T_2\in\mathscr{U}$（注意$\mathscr{U}$虽被假定为非超滤集，但$\mathscr{U}$至少是一个滤集，故由性质2°知此为真）。而$X_1\cap X_2\supseteq S\cap T_1\cap T_2$，故按定义$X_1\cap X_2\in\mathscr{U}'$。

可见，$\mathscr{U}'$确实是N上的一个滤集。另一方面，对每一$G\in\mathscr{U}$，恒有$G\supseteq S\cap G$，故$G\in\mathscr{U}'$。因此$\mathscr{U}\subseteq\mathscr{U}'$。又$S\supseteq S\cap T$，故$S\in\mathscr{U}'$，但$S\notin\mathscr{U}$。这表明$\mathscr{U}'$确实包含$\mathscr{U}$但又大于$\mathscr{U}$。这与$\mathscr{U}$为极大元的性质矛盾。由此证明上述$S$与$\mathscr{U}$每一元的交集**均非空**是不能成立的。因此可以断言：

(1) 至少存在一个N的子集$T\in\mathscr{U}$致使$S\cap T=\varnothing$。完全类似地可以推证：

(2) 至少存在N的子集$T'\in\mathscr{U}$使$(N-S)\cap T'=\varnothing$。由(1)(2)易见$T\subseteq(N-S)$，$T'\subseteq S$，$T\cap T'=\varnothing\in\mathscr{U}$。但这最后一式与$\mathscr{U}$的滤集性质1°相矛盾。

故上述$\mathscr{U}$只能是N上的一个超滤集。定理证毕。

（三）超幂$^*R$的构造. 利用超滤集$\mathscr{U}$可以对R^N中的函数界定

各种属性而不致违反逻辑上的排中律。例如，设P代表某一性质，$\{f(n)\}\in R^N$，则称$f=\{f(n)\}$具有性质P，意即

$$\{n\in N\,|\,f(n)\text{具有性质}P\}\in\mathscr{U}.$$

简言之，这就是按照$\mathscr{U}$来界定“f具有性质P”这一概念。我们还可按照$\mathscr{U}$来界定两个函数间的等价概念。回到我们原来的出发点，仍记$x_n=f(n)$，$y_n=g(n)$，等等，这里f，$g\in R^N$。

等价概念 设$\tilde{x}$与$\tilde{y}$是两个过程量，且$\tilde{x}\longleftrightarrow\{x_n\}$，$\tilde{y}\longleftrightarrow\{y_n\}$。又设$\{n\,|\,x_n=y_n\}\in\mathscr{U}$。则称$\{x_n\}$与$\{y_n\}$等价，或称相应的过程量$\tilde{x}$与$\tilde{y}$彼此等价，记作$\tilde{x}=\tilde{y}$。

显然如上定义的等价关系“$=$”具有反身性、对称性与传递性。这样就可把R^N中的一切数列$\{x_n\}$按等价关系来分类。例如，同$\{x_n\}$等价的全体数列构成的类可以记作$\langle x\rangle$。注意$\{x_n\}\longleftrightarrow\tilde{x}$，因此，$\langle x\rangle$同时也可看成为$\tilde{x}$的等价类。

于是，所谓超幂$^*R$就是由一切过程量等价类$\langle x\rangle$作成的集合，即$^*R=\{\langle x\rangle\}$。

值得注意的是，由于一个等价类$\langle x\rangle$完全由一个过程量$\tilde{x}$或相应的阶梯函数$x(t)$所唯一确定，所以，对任何一个等价类$\langle x\rangle$可以把相应的$\tilde{x}$或$x(t)$当作$\langle x\rangle$的代表。特别，对应于常过程量$\tilde{a}$、$\tilde{0}$，$\tilde{1}$的等价类可分别记作$\langle a\rangle$，$\langle 0\rangle$，$\langle 1\rangle$等等，它们又可简单地用$a,0,1$等等来代表。这样看来，由所有常量c构成的实数集$R=\{c\}$可以看成是嵌入在$^*R$内的一个子集。

（四）三分律的验证．可以验明$^*R$是一个满足三分律的有序集。对于$^*R$中的$\langle x\rangle$与$\langle y\rangle$，当

$$\{n\in N\,|\,x_n\leqslant y_n\}\in\mathscr{U}$$

时，则定义$\langle x\rangle\leqslant\langle y\rangle$。这个定义是合理的，因为如果$\langle x\rangle=\langle x'\rangle$，$\langle y\rangle=\langle y'\rangle$，则易证

$$\langle x\rangle\leqslant\langle y\rangle\Longleftrightarrow\langle x'\rangle\leqslant\langle y'\rangle.$$

事实上，由

$$F=\{n\in N\,|\,x_n=x_n'\}\in\mathscr{U},$$

$$G=\{n\in N \mid y_n=y'_n\}\in\mathscr{U},$$
$$H=\{n\in N \mid x_n\leqslant y_n\}\in\mathscr{U},$$
易见$F\cap G\cap H\subseteq\{n\in N \mid x'_n\leqslant y'_n\}$，$F\cap G\cap H\in\mathscr{U}$，故
$$\{n\in N \mid x'_n\leqslant y'_n\}\in\mathscr{U}.$$
这即表明$\langle x'\rangle\leqslant\langle y'\rangle$，反之可由$\langle x'\rangle\leqslant\langle y'\rangle$推证$\langle x\rangle\leqslant\langle y\rangle$。

所谓$\langle x\rangle\neq\langle y\rangle$意即
$$J=\{n\in N \mid x_n\neq y_n\}\in\mathscr{U}.$$
如果$\langle x\rangle\leqslant\langle y\rangle$且$\langle x\rangle\neq\langle y\rangle$，则$H\cap J\in\mathscr{U}$，从而
$$H\cap J\subseteq\{n\in N \mid x_n<y_n\}\in\mathscr{U},$$
故此时可记$\langle x\rangle<\langle y\rangle$ 。现在我们来证三分律。

三分律 对任何$\langle x\rangle$，$\langle y\rangle\in{}^*R$，在下列三式中
$$\langle x\rangle=\langle y\rangle,\quad \langle x\rangle<\langle y\rangle,\quad \langle y\rangle<\langle x\rangle$$
有且仅有一式成立。

证明 令
$$E=\{n\in N \mid x_n=y_n\},$$
$$L=\{n\in N \mid x_n<y_n\},$$
$$G=\{n\in N \mid y_n<x_n\},$$
只要证明E，L，G中有且仅有一个在$\mathscr{U}$中就可以了。首先，让我们注意到在标准数域R中三分律是成立的，即关系式$x_n=y_n$，$x_n<y_n$，$y_n<x_n$三者中恰有一式成立。因此$E\cup L\cup G=N$，并且
$$E\cap L=E\cap G=L\cap G=\varnothing \text{（空集）}.$$
以下先证E，L，G中至少有一个是属于$\mathscr{U}$的自然数子集。倘若$E\in\mathscr{U}$，则结论已真。否则如果$E\notin\mathscr{U}$，则因$\mathscr{U}$为超滤集，故$N-E=L\cup G\in\mathscr{U}$.此时若$L\in\mathscr{U}$，则结论已真，否则$L\notin\mathscr{U}$，则$N-L=E\cup G\in\mathscr{U}$。于是由滤集性质2°知
$$(L\cup G)\cap(E\cup G)=G\in\mathscr{U}.$$
这表明E，L，G中至少有一个属于$\mathscr{U}$。另一方面，可证E，L，G中不能有两个同时在$\mathscr{U}$中。否则，由滤集性质2°，它们的交集要在$\mathscr{U}$中.上面已知它们的交集为$\varnothing$，于是$\varnothing\in\mathscr{U}$.这就矛盾于滤集

性质1°. 故结论是E, L, G中只有一个在$\mathscr{U}$中。证毕。

有了三分律，我们便可进一步把有序集$^*R$做成一个具有三分律的**有序域**。这只要在$^*R$内引进四则运算即可。

（五）作为有序域的$^*R$. 在过程量之间我们已经规定了加、减、乘等运算。请再注意，$\langle x \rangle$，$\langle y \rangle$等无非是$\langle \tilde{x} \rangle$，$\langle \tilde{y} \rangle$等缩记。于是$^*R$中元素间的运算可定义如下：

$$\langle x \rangle \pm \langle y \rangle = \langle \tilde{x} \rangle \pm \langle \tilde{y} \rangle = \langle \tilde{x} \pm \tilde{y} \rangle = \langle x \pm y \rangle,$$

$$\langle x \rangle \cdot \langle y \rangle = \langle \tilde{x} \rangle \cdot \langle \tilde{y} \rangle = \langle \tilde{x} \cdot \tilde{y} \rangle = \langle x \cdot y \rangle .$$

易知，在这样定义之下的加法和乘法分别满足交换律、结合律，并且加法与乘法适合分配律。又$\langle 0 \rangle$，$\langle 1 \rangle$和$\langle -x \rangle$分别是加法零元素、乘法单位元素和$\langle x \rangle$的加法逆元素。

乘法逆元的定义稍稍麻烦一点。今设$\langle x \rangle \neq \langle 0 \rangle$，即

$$\{n \in N \mid x_n \neq 0\} \in \mathscr{U}.$$

于是把过程量$\tilde{x}$的倒数$\tilde{y} = 1/\tilde{x} = \tilde{x}^{-1}$引出的等价类$\langle \tilde{y} \rangle = \langle \tilde{x}^{-1} \rangle = \langle x^{-1} \rangle$规定为$\langle x \rangle$的逆元，即规定$\langle x \rangle^{-1} = \langle y \rangle = \langle x^{-1} \rangle$。由此，可见

$$\langle x \rangle \cdot \langle x \rangle^{-1} = \langle x \rangle \cdot \langle y \rangle = \langle x \cdot y \rangle = \langle \tilde{x} \cdot \tilde{x}^{-1} \rangle = \langle 1 \rangle .$$

事实上，最后一式的导出是因为

$$\{n \in N \mid x_n \cdot x_n^{-1} = 1\} \in \mathscr{U}.$$

综上所述，可知$^*R$是一个具有四则运算并满足三分律的有序域，而标准实数域R自然是它的一个子域。

（六）非标准实数. 仿照实数的绝对值概念，可以对$^*R$中的元素引进绝对值概念：

$$\left| \langle x \rangle \right| = \begin{cases} \langle x \rangle & \text{当} \langle x \rangle \geqslant \langle 0 \rangle \text{时}, \\ -\langle x \rangle & \text{当} \langle x \rangle < \langle 0 \rangle \text{时}. \end{cases}$$

$|\langle x \rangle|$可简记为$|x|$。今引进下述定义。

定义 设$x \in {}^*R$（即$\langle x \rangle \in {}^*R$），如果对任意的正实数$r$，恒有$|x| < r$，则称$x$为无限小。又如果对任意的正实数$r$，恒有$|x| > r$，则称$x$为无限大。假如存在一个固定正实数$r$使得$|x| \leqslant r$，则称

x为有限数。凡无限小、无限大及有限数与无限小之和都称为非**标准实数**。

容易验明，前面定义过的趋零与趋无穷过程量的等价类就是符合上述定义的无限小与无限大。由于这些等价类多至无穷，故$^*R$中确实存在着无穷多的无限小与无限大。

（七）$^*R$中数的结构及非标准数轴。我们首先指出，在普通数学分析里存在这样的变量：它既不是有界量也不是无限大。但在非标准数域$^*R$里这种情况是不存在的，即对任何$x\in {}^*R$，要么x是无限大，要么x是有限数。这是因为按照超滤集$\mathscr{U}$来判定$^*R$中成员的性质时，不等式关系$|x|\leqslant r$与$|x|>r$（r为任意正实数）是不能同时成立的。从这里还可以看到引用超滤集概念的必要性：利用超滤集作为判定各种性质的真假标准时，正可以排除模棱两可的情形。事实上，如果不引入超滤集作为判定成员大小的标准，则$^*R$集合的有序性和三分律即无法建立起来。

由上所述，我们可将$^*R$中的数分成两类，一类是有限数，一类是无限大。对于有限数，我们有如下的定理。

结构定理　$^*R$中的每个有限数α，皆可唯一地表示成这样的形式

$$\alpha=\text{标准实数}+\text{无限小}。$$

此处所谓标准实数是指子域R中的实数（过程化了的），并且0（即$\langle 0\rangle$）也是一个无限小。

证明　先证表示法的存在性。设$\alpha\in {}^*R$是有限数，我们要证明$\alpha=r+\varepsilon$，其中r是标准实数，ε是无限小。为此，试考虑集合

$$A=\{x\,|\,x<\alpha，且x\in R\}，$$

既然α是有限数，至少存在$r_0\in R$，使$\alpha<r_0$，故A是有界集合。又$A\subset R$，故有上确界$a=\sup A$。于是$|\alpha-a|$或为实数，或者不是实数（标准实数）。

（1）若$|\alpha-a|$是实数，不妨记$|\alpha-a|=h\in R$，于是$\alpha-a=\pm h$，

故$\alpha=(a\pm h)+0$，0也是无限小，故定理得证。

(2) 若$|\alpha-a|$不是实数，则可证$\alpha-a$是无限小。今用反证法证之。设$|\alpha-a|$不是无限小，则必存在某个正实数$h\in R$使$|\alpha-a|>h$。下面分两种情形考虑：

(i) $\alpha-a>h$，这时$\alpha>a+h$，故$a+h\in A$。但a是A的上确界，故$a\geqslant a+h$，于是$0\geqslant h$，矛盾。

(ii) $a-\alpha>h$，这时$a-h>\alpha$，这与a是A的上确界相矛盾。

因此$\alpha-a$是无限小。于是由$\alpha=a+(\alpha-a)$便得到所要求的表达式。最后用反证法还可证定理中的表达式是唯一的。至此定理得证。

由此可见，$^*R$中的数是由标准实数，无限小和无限大组成的。有限数α表示形式中的标准实数叫作α的**标准部分**(Standard part)，记作$^\circ\alpha$（特别，α本身为标准实数时，则$^\circ\alpha=\alpha$）。对$^*R$中的任意有限数α，β等我们有下列简单等式

$$^\circ(\alpha+\beta)={}^\circ\alpha+{}^\circ\beta,\quad {}^\circ(\alpha-\beta)={}^\circ\alpha-{}^\circ\beta,$$

$$^\circ(\alpha\cdot\beta)={}^\circ\alpha\cdot{}^\circ\beta,\quad {}^\circ\left(\frac{\alpha}{\beta}\right)=\frac{^\circ\alpha}{^\circ\beta}\ (\beta\text{不是无限小时})。$$

如所知，在古代曾误认有理数与数轴上的点是一一对应的，后来发现了无理数，人们又长期地把实数看成和数轴上的点是一一对应的。这是基于这样一条公理，即点是不可分的，是没有内部结构的。现在又发现了（实际是界定了）无限小和无限大这类非标准数，它们在数轴上应赋予其位置。但是，不可分的实数点已经充满了实数轴，因此，要想在实数轴上找到无限小和无限大这类非标准数的位置就得设想（或假定）：第一，实数点是可分的，也就是说，在实数点的“内部”凝聚着数不清的无限小点，第二，在数轴的两端（无限远处）还凝聚着数不清的无限大坐标点。当我们把数轴作了这样理解后，就称为非标准数轴。如此，非标准数轴上的点和$^*R$中的数可以看作是一一对应的。这就是非标准数域$^*R$的几何解释。

建立了非标准有序域$^*R$之后，在其上再建立非标准微积分学就不再遇到任何困难了。首先，我们可以在$^*R$上对函数$f(x)\in(^*R\to{}^*R)$界定微商概念。令ε为无穷小，设$x_0\in{}^*R$。如果下述标准部分存在：

$$f'(x_0)={}^\circ\left(\frac{f(x_0+\varepsilon)-f(x_0)}{\varepsilon}\right),$$

则称标准实数$f'(x_0)$为$f(x)$在x_0处的微商。

至少连续函数的定积分概念也不难描述如下：设$g(x)$是$^*R\cap[\alpha,\beta]$上的连续函数。令ω为无限大自然数，设π表示$[\alpha,\beta]$的无穷小分划

$$\pi:\ \alpha=x_0<x_1<x_2<\cdots<x_\omega=\beta,$$

其中$x_k=\alpha+k\varepsilon$，$\varepsilon=(\beta-\alpha)/\omega$，$(k=0,1,\cdots,\omega)$。于是$g(x)$在区间$[\alpha,\beta]$上的定积分可表示为下述标准部分：

$$\int_\alpha^\beta g(x)dx:={}^\circ\left(\sum_{k=1}^{\omega}g(\xi_k)\cdot\varepsilon\right),$$

这里$x_{k-1}\leqslant\xi\leqslant x_k$，而关于$\{\xi_k\}$的任意取法均不影响上述标准实数（积分值）的一意确定性。验证甚易，此处从略。有关细节请阅读专书。

§3 非康托型自然数序列模型的构造法

利用过程量概念可以很容易地给出$^*R$中的无限自然数概念。设$x_n(n=0,1,2,\cdots)$是N^N中的数列，即$\{x_n\}$是一个恒取非负整数值的序列。如果$x_n\uparrow+\infty(n\to\infty)$，则$\{x_n\}$相应的过程量$\tilde{x}$的等价类$\langle x\rangle=\langle\tilde{x}\rangle$就叫作$^*R$中的无限大自然数。显然这个定义是合理的，因为$\tilde{x}$所对应的整点阶梯函数恒取整数值，且易验明对任给标准自然数$\langle m\rangle$（常量过程的等价类）都有$\langle x\rangle>\langle m\rangle$。

例如，令$[a]$表示标准正实数a的整数部，则易见

$$x_n = n,\ x_n = \left[\frac{n}{2}\right],\ x_n = [\sqrt{n}],\ x_n = n^2 - 5$$

等相应地导出的等价类 $\langle x\rangle$ 都是无限大自然数，它们可以分别地简记作

$$\omega,\ \left[\frac{\omega}{2}\right],\ [\sqrt{\omega}],\ \omega^2 - 5,\ 等。$$

在康托的序数理论里，把ω视为诸n无限增大时趋向的目标，所谓极限序数，或者把ω定义为自然数序列的序型。基于这样一种看法，则自然数序列的有限平移，例如将$1, 2, 3, \cdots, n, \cdots$改成为$k+1,\ k+2,\ \cdots,\ k+n,\ \cdots$或$-k+1,\ -k+2,\ \cdots,\ -k+n,\ \cdots$之后，其无限增大时的趋向目标完全雷同，无限增大时的步调完全一致（均为逐步增大）。这就启发我们可以把$^*R$中的无限自然数$\omega \pm k$（k为任一固定自然数）与ω视为同一类对象而汇集成一个类。详见如下定义：

定义1 称形如$\omega \pm k$的无限大自然数为ω的等价元，而由所有等价元汇集成的类称为ω的等价类或ω银河（galaxy），简记作（ω）。如果无限大自然数ν比（ω）中的每个成员$\omega \pm k$都小，则称ν小于(ω)，并可将此种关系表示为$\nu < (\omega)$。

作为举例，如取$\nu = \left[\frac{\omega}{2}\right]$，$\nu = [\sqrt{\omega}]$，等，则它们都比（$\omega$）为小。（验证极易，此处从略）。

容易想到，在形式上我们可以作出如下对应关系：

$$\begin{array}{ccccc} 1, & 2, & 3, \cdots, & n, \cdots, & \omega \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow \\ \langle 1\rangle, & \langle 2\rangle, & \langle 3\rangle, \cdots, & \langle n\rangle, \cdots, & \langle\omega\rangle \end{array}$$

但是我们的目的是要构造出一个**非康托意义的**自然数序列模型，因此需要对§1中所谈论到的飞跃段赋予具体含义（当然没有必要去考虑比ω或（ω）大得多的无限自然数，如ω^2，ω^3，…等等）。让我们引进下述定义。

定义2 令$\langle n\rangle$表示$^*R$中任意有限自然数。则所有小于(ω)的无限大自然数ν按$^*R$中的序关系构成的序集片段叫作$\langle n\rangle$与

(ω) 之间的飞跃段（或**潜变段**），可简记作

$$\widetilde{N}=\{\nu \mid \nu<(\omega),\ \nu\text{为无限自然数}\}.$$

如上所述，在$^*R$中飞跃段$\widetilde{N}$这个概念是十分具体而明确的，但把$^*R$中的一切$\langle n\rangle$考虑成为从延伸到穷竭的过程时，该过程是否包含N的问题，只能依靠公理来规定。为此，我们引进如下公理：

飞跃段存在公理 在$^*R$中经由延伸到穷竭而形成的正整数序列过程

$$\langle 1\rangle, \langle 2\rangle, \langle 3\rangle, \cdots, \langle n\rangle, \cdots$$

必将穷尽并包含了比(ω)小的所有标准自然数与无限大自然数。从而在上列过程中必定包含着飞跃段$\widetilde{N}$。

既然R中的n与$^*R$中的$\langle n\rangle$作成形式上的一一对应，且康托序数ω与(ω)也可对应起来，则飞跃段存在公理对应地应用于通常的自然数序列过程时，将同样可以引出一个飞跃段概念，当然它已经不再属于康托意义的自然数序列范畴。由于这个缘故，我们理应把包含着飞跃段$\widetilde{N}$的$^*R$中的自然数序列模型（或其和标准自然数序列过程的对应物）称之为**非康托自然数序列模型**（*Non-Cantorian model of natural number sequence*）。以后记这种模型为$\mathbf{N}=\{\nu \mid \nu<(\omega)\}$。

如果把(ω)称之为真实无限大，则$\widetilde{N}$中的各个无限大可以叫作过渡性的无限大或潜变无限大。这样，$\widetilde{N}$称之为真实无限过程中的潜变段也是很有理由的。恩格斯关于无限性的悖论断言：无限是由有限发展而成的。现在用我们的非康托模型来解释就很清楚了，有限和无限以及潜无限和实无限的相互渗透是通过潜变段$\widetilde{N}$来实现的。潜变段（即$\widetilde{N}$）是由无穷多个潜变无限大元素组成的序集片段，它是贯通从有限到真无限的桥梁。另一方面，康托意义下的标准自然数序列模型$N=\{1, 2, 3, \cdots, n, \cdots\}(n<\omega)$，恒假定序列过程中的成员的有限性是守恒不变的，这样就无法理解具有无限性的真实无限过程是如何产生出来的。因此这种经典

性的自然数序列模型是不能很好地用来解释恩格斯无限性悖论的实质的。

还值得注意的是，非康托模型$N=\{\nu|\nu<(\omega)\}$中的元素个数已经不是可数无限多。事实上，易证下述命题为真：

命题 在非康托自然数序列模型N中至少包含着同实数一般多的无限大自然数。

证明 在$^*R$上进行作业。试取无限自然数列

$$\left[\frac{\omega}{3}\right],\ \left[\frac{\omega}{3^2}\right],\ \left[\frac{\omega}{3^3}\right],\ \cdots,\ \left[\frac{\omega}{3^n}\right],\ \cdots.$$

然后作形如下状的无限大自然数

$$\lambda_1\left[\frac{\omega}{3}\right]+\lambda_2\left[\frac{\omega}{3^2}\right]+\lambda_3\left[\frac{\omega}{3^3}\right]+\cdots+\lambda_n\left[\frac{\omega}{3^n}\right]+\cdots,$$

其中诸λ_i取0或1（但$\lambda_1=1$），故上述无限大自然数的个数同二进位小数一般多。按非标准分析算法容易验明

$$\left[\frac{\omega}{3}\right]+\left[\frac{\omega}{3^2}\right]+\left[\frac{\omega}{3^3}\right]+\cdots+\left[\frac{\omega}{3^n}\right]+\cdots$$

$$\leqslant\frac{\omega}{2}<\omega\pm k,$$

其中k为任一有限正整数。命题证毕。

最后，让我们指出两点：（一）N中的$\widetilde{N}$与$^*R$中的非标准自然数集合$^*N$是不同的。首先，两者在内容上是不同的。因为在$^*N$中包含有2ω，3ω，ω^2，$\omega-1$，$\omega-2$，…，但在N中的无限大自然数都比形如$\omega\pm k$的数为小。其次，$^*N$中的非标准自然数是通过外插法引进的，而N中的无限自然数固然是由过程量等价类来定义的，但一旦定义之后即可理解为它们是同无穷总体$N=\{\nu|\nu(\omega)\}$的生成过程（即自然数(n)从延伸到穷竭的过程）不可分割的，因为$\widetilde{N}$正好是构成N真无限性的关键部分。总之，根本区别乃在于：N及$\widetilde{N}$乃是被解释成**生成观点下的**无限性总体，而$^*N$却是在**外插观点下**由N扩引而成的。（二）N中的无限自然数既无最小者也无最大者，这就是我们所以把$\widetilde{N}$称之为(n)

与 (ω) 间的潜变段或飞跃段的理由所在。根据我们证明了的命题，固然$\mathbf{N}$及其片段$\widetilde{N}$都是有序集合，但其中的元素却是无法按可数无限次步骤取尽的。

附记 按飞跃段存在公理，在$^*\mathbf{R}$上考虑一切$\langle n\rangle$时则得$\mathbf{N}$. 但在$\mathbf{R}$上考虑一切n时则得到标准自然数集合N. 因此"一切"（*all*）有两种不同含义，在不同的论域上来运用"一切"这一全称量词时，可以得出不同的内容. 既然对初始片段而言，存在一一对应$\langle n\rangle \longleftrightarrow n$，所以在记法上不妨把$\langle n\rangle$简记为$n$. 但为了区别$N$和$\mathbf{N}$，可分别采用下述序列表示方式：

N：1，2，3，…，n，…

$\mathbf{N}$：1，2，3，…，n，…

或相应地简记为$\{n\}$与$\{\langle n\rangle\}$. 这样就不致于彼此混淆了.

§4 关于一个引伸的芝诺悖论的解释

在上节中我们给出了飞跃段$\widetilde{N}$的内部结构和非康托自然数序列模型$\mathbf{N}$。现在我们利用这种模型可以对一个"引伸的芝诺悖论"（Extended Zeno's paradox）给出一种异乎寻常的解释。

引伸的芝诺悖论又叫作**抛球问题**。设有A，B二人抛一球，球在A，B二人手中抛来抛去。假设经过的时间（分钟）依次为$\frac{1}{2}$分钟，$\frac{1}{4}$分钟，$\frac{1}{8}$分钟，…，$\frac{1}{2^n}$分钟（$n=1,2,3,\cdots$），而球将依次到达A和B处。试问总时间到达一分钟时，该球到达何处？

记
$$S_n=\sum_{k=1}^{n}\frac{1}{2^k},\quad (n=1,2,\cdots),$$

则
$$\lim_{n\to\infty} S_{2n}=\lim_{n\to\infty} S_{2n+1}=1.$$

显然在各个时刻S_{2n}，球总是落在B处；而在各个时刻S_{2n+1}（$n=1,2,\cdots$）球总落在A处。但时间序列$\{S_{2n}\}$与$\{S_{2n+1}\}$的极限都是1，故到达一分钟时就无法判断该球究竟落在何处。说它落在A

处或B处都有“理由”但又都会引出矛盾。当然这样一个涉及时间无限分割过程的悖论在现实生活中是不存在的，但在理性思维领域却有分析思考的必要。下面将指出这种悖论在非标准数域上的**相对存在性**。

我们来分析时间序列$\{S_n\}$。如果在时间区间$\frac{1}{2}\leqslant S\leqslant 1$内的时间分划数$n$规定只走遍康托意义下的标准自然数序列$N=\{1, 2, 3, \cdots\}$，则上述悖论即不复存在。现在我们假定$n$走遍N中的一切有限与无限自然数。令

$$\nu_k=(2k)\left[\frac{\omega}{2k+1}\right],\ \mu_k=(2k)\left[\frac{\omega}{2k+1}\right]+1 \quad (k=1,2,3,\cdots).$$

则ν_k与μ_k分别为无限大偶自然数与奇自然数，且容易验明$\nu_k<(\omega)$，$\mu_k<(\omega)$，故ν_k，$\mu_k\in$N。已知于时刻S_{2n}与S_{2n+1}时球将轮流落到B处与A处。不妨认定对以无限大偶数与奇数为足标的时刻S_{ν_k}与S_{μ_k}而言，上述情况仍继续保持。这样，由于按非标准分析方法我们有

$$°S_{\nu_k}=1,\ °S_{\mu_k}=1,\quad (k\in N).$$

可见于标准时刻1分钟时，该球将无限多次地落在A处，同时又将无限多次地落在B处。这样，就形成一个悖论。

上述结论看来非常奇怪。但只要想到时点如同非标准数轴上的实数点那样，它是具有内部结构的，那就不足为奇了。事实上，我们正是利用了非标准实数（例如S_{ν_k}与S_{μ_k}）的这种能够无限多地凝聚在标准实数（例如1）无限小邻域内的特点。

§5　略论无限的两种形态

对自然数序列无限性的分析，使我们看到了无限性的两种形态或两种结构。康托意义的自然数序列的无限性是一种形态；非康托自然数模型提供了另一种形态。前一形态称为**标准形态**，后

一形态称为**非标准形态**。

标准形态的无限性是靠ω来表征的，ω代表着整个标准自然数序列"1,2,3,…, n, …"的真无限性，但序列中的序数具有守恒的有限性，只有有限性的变量，而无转化为无限性的飞跃或质变。这种形态的无限性与ω之前的有限性是截然分裂的，故符合形式逻辑的排中律：非有限即无限。即真无限性乃是有限性的直接否定。

ω是在整个自然数序列过程完成时引出的。在ω引出之前并不存在任何中间状态或潜变形式的无限性。因此，这里实际上已经用到了一条基本公理，即如下所述的**无限凝聚公设**：具有真无限性的ω是整个自然数序列完成时的产物，而序列的过程中不包含任何形式的无限性（连潜无限性也不存在）。

非标准形态的无限性是由**N**的飞跃段（潜变段）$\widetilde{N}$来表现的。在这里我们用到了另一条基本公理，即如下所述的**无限发散公设**：具有真无限性的（ω）虽是整个自然数序列完成时的产物，但在序列过程中已经包含有一个飞跃段$\widetilde{N}$，飞跃段中有着不可数无穷多的潜变无限大元素（较（ω）为小的无限大自然数），它们组成了（ω）的单侧无限大邻域。

由上所述，可见标准形态的无限性可以称为**凝聚态的无限性**，而非标准形态的无限性又可称为"发散态的无限性"。前者由ω集中表现，后者由$\widetilde{N}$与（ω）的联合来表现。

显然上面关于无限性的"凝聚公设"与"发散公设"是彼此矛盾的。它们正好比欧氏几何的平行公理（第五公设）与罗氏几何平行公理相互矛盾一样。但是由于非标准形态的无限性来源于非康托自然数序列模型**N**的确立，而**N**的确立是由于非标准数域$^*R$的存在性才得到保证的。我们又知道$^*R$的构成，归根结底是利用了二维的实数空间$R\times R$上的标准分析学方法（当然还用到了与选择公理相等价的Zorn引理）。因此，根据希尔伯特特化归法思想可知，只要标准实数域R是无矛盾的，则$^*R$即不应有矛

盾，从而N模型的存在性（在非标准分析中）也不应有矛盾。

但R上的分析学有无矛盾，可归结为N上的算术有无矛盾。因此归根结底只要假定标准自然数系统无矛盾，也就是以保证非标准自然数模型N的构造方式的合理性（即在非标准分析领域中的合理性）。

以上的解释无非说明**“无限发散公设”**的相对合理性。因此它正好象罗巴契夫斯基几何学中的（罗氏）平行公理那样，完全可以根据该公设，自行演绎出一套完全不同于康托经典理论的新理论。新的理论系统应该是没有矛盾的，如果经典的理论系统是没有矛盾的话。

但是，将无限发散公设或发散态的无限性概念推广到高阶超限序数上去，看来兴趣是不大的。倒是模型N本身（或飞跃段$\widetilde{N}$的结构特征）及其含义解释给经典逻辑法则带来的麻烦值得注意。这里主要指的是排中律问题。本来，象有限与无限两个概念是泾渭分明的，但在模型N中却有着连系有限与无限（ω）的潜变段$\widetilde{N}$，这样就有着**既非有限又非真无限**的潜变无限大，它们的存在正好背离了排中律。上述潜变段或飞跃段乃是自然数序列过程从延伸到穷竭的生成物，所以$\widetilde{N}$又是无法从整个序列过程中分离出去的。

由上所述，如果按N模型来思维自然数无限性的话，则由于排中律的失灵，我们在数学的分析论证推理上，就势必要被迫接受近代直觉主义者的观点：即当**论域是无限对象**时，须采用**直接论证法或构造法**以代替反证法（或归谬法）。这样，可以说我们和直觉者的主张殊途同归了。当然两者的出发点却是完全不同的。

两种形态的无限性概念的深入讨论，必然很费篇幅，值得大作文章。这里只好从略了。

附记　笔者在1980年发表的一篇题为“略论近代数学流派的无限观和方法论”一文中，曾以自然数序列无限性为例，述及双相无限（double -phased infinity）

的概念，即无限过程乃是真无限与潜无限的对立统一体，两种无限概念只不过是对同一个无限性对象（如自然数序列）的两个侧面的模写或反映．只要略作深入分析即可发现，数学上的每一种无限过程，本质上都是双相无限．但由于形式概念思维的单一式和僵化性，在考察无限性的任一侧面时，另一侧面必处于形式推理上被否定的地位．所以，在肯定关于实无限的凝聚公设时，必定要否定潜无限的存在性；另一方面，潜无限论者就必然否定真无限的存在性，否则将自相矛盾．这样说来，倒是无限发散公设所肯定的非康托模型N能将彼此矛盾着的两个侧面——潜无限与真无限联结成一个对立统一体(所以N乃是一个表现双相无限的模型）．但这样一来，基本上以有限论域为背景抽象出来的逻辑上的排中律就不能无条件地适用了．

近代直觉主义者是以坚执无限性的单一侧面（潜无限观点）来否定无限论域上的排中律的，而我们乃是以表现双相无限的模N型的存在性为依据来得出同样结论的．关于双相无限概念在数理哲学上的意义，曾在1981年郑毓信同志的硕士学位论文《论数理哲学中的几个问题》中作过较深入的分析．

第8讲　悖论与数学基础问题

§1　悖论的定义和起源

所谓悖论，从字面上讲就是荒谬的理论。那么，为什么要把悖论这样一个晦涩古怪的名词来取代这一通俗易懂的说法呢？按照美国柯朗数学研究所M. Kline教授的说法，那是为了不把自相矛盾的真相摆在桌面上，才采用了这样一个婉转的措辞。当然，M. Kline教授是针对Paradox才这样说的，并不是针对“悖论”所说，因此，我们在这里不是直接引用而是引而套用了M. Kline的说法。当然，现在时间一长，悖论这一词汇也用习惯了，其真实函义也就人人皆知了。

关于悖论的定义，当前流行的说法，如悖论是一种导致逻辑矛盾的命题。这种命题，如果承认它是真的，那么它又是假的；如果承认它是假的，那么它又是真的。又如悖论是指这样一个命题A，由A出发，可以找到一语句B，然后，若假定B真，就可推得$\neg B$真，亦即可推出B假。若假定$\neg B$真，即B假，又可推导出B真。再有如一个命题构成一个悖论，如果由它的真可以推出它的假，而由它的假又可以推出它的真。诸如此类的定义法，有它合理的一面，但有不够全面的地方。第一，任何一个悖论在实质上都相对地被包含在某个理论体系中，例如，著名的罗素悖论是一个被包含在古典集合论系统中的悖论。可能有的悖论看上去似乎不象针对那一个理论体系，其实只是这个悖论所属系统的原始概念和基本原则没有明朗化。所以，一个悖论总是相对于某个理论系统而言的。如果竟有这样一个悖论，它将被包含在历史的和将来的任何一个理论体系中，那么，我们就既不要去研究什么悖

论的成因，也不必去考虑排除悖论的方法，因为既有这种绝对的悖论出现，那么研究它也无济于事。当然，我们也不相信会有这种悖论出现。由此可见，当我们给悖论下定义的时候，如果忘记了相对于某一理论体系这个前提，将会造成怎样的误解。第二，并非每个悖论都要陈述为一个命题或某一语句的形式，有的悖论往往要有一个推演过程来表现。第三，从一些著名的悖论的陈述形式来看，人们并不习惯于要求把每个悖论都化归为肯定等价于否定的形式，也可用某一系统中并存的两个互相矛盾的命题来表述一个悖论。例如，古典集合论中的康托悖论就是一例，人们习惯于把它表述为$\overline{\overline{M}}<\overline{\overline{P(M)}}$和$\overline{\overline{M}}\geqslant\overline{\overline{P(M)}}$这样两个互相矛盾的命题，并没有也不必要化归为两个矛盾命题的等价式。综上所述，我们认为孤立地用肯定等价于否定来作为悖论的定义是不够合理而全面的。

所以，我们主张采用 A.A.Fraenkel 与Y.Bar-Hillel 的说法，如果某一理论的公理和推理原则看上去是合理的，但在这个理论中却推出了两个互相矛盾的命题，或者证明了这样一个复合命题，它表现为两个互相矛盾的命题的等价式。那么，我们就说这个理论包含了一个悖论。我们认为，这样来定义悖论较为全面而合理，因为在这里，首先指明了任何一个悖论总是相对于某一理论系统而言的。其次又指出了一个悖论可以表现为某一理论系统中两个互相矛盾的命题的形式。最后才指出，悖论也可集中地表现为肯定等价于否定的复合命题。这样看来，当前流行的那种定义法，就只是抽取了 Fraenkel 陈述中的最后两句话作为悖论的定义，这当然是不够全面而合理了。另外，在 Fraenkel 陈述中的第一句话，不光是指明了任一悖论总相对于某一系统这一点，十分重要的是在这里强调了该系统的公理和推理原则看上去是合理的，如果不强调指出这一点，那么我们就可轻而易举地把一些明显的矛盾命题凑合在一起算作一个系统，然后宣布在这个系统中创造性地发现了许多悖论。

关于悖论的起源，可以追溯到古希腊和我国先秦哲学时代，但在那时及其往后的一个相当长的历史时期中，悖论往往泛指那些推理过程看上去是合理的，但推理的结果却又违背客观实际。例如，著名的芝诺悖论便属于这一类悖论。

在历史上，还有另一种与之相反的情形而称之为悖论的，那就是由于新概念的引入而违背了具有历史局限性的传统观念，这也一时称为悖论的发现。这就不是推理看上去好象是合理的问题，而是传统观念貌似真实的事了。例如Galilei对平方数与自然数一一对应的发现而矛盾于全体大于部分的原则，在历史上也称为Galilei悖论。这就不是Galilei的发现在推理上的问题，而是由于全体大于部分的直观原则是从有限数量的事物关系中抽象出来的，自然不适用无穷集合的情形了。诸如此类的悖论还可列举，例如，鳄鱼的二难论等，但所有这些与我们今天所讲悖论的含义已略有距离。因之，述其一、二就不作更多的讨论了。

历史上，与今天所讲悖论的含义较为切近并可看作悖论之直接起源的是这样一件事。公元前六世纪，克里特哲学家Epimenides发现了一个实际上没有构成悖论的悖论。其原始命题是：一个克里特人说："所有的克里特人所说的每一句话都是谎话"。试问这句话是真是假。如果它是真话，则因这句话也出自一个克里特人之口，故按此话的论断可推知这句话为假。因之，由这句话的真可导至它为假。但反过来，若设这句话为假，则并不导致任何矛盾。但仅由它的真可导致它为假这一点而言，就足以引人注目了。

在此顺便指出：上述那个并非悖论的悖论却至今有被误认为悖论的情况，其实是一种误解或疏忽。因为假定这句话是假，并不引起矛盾，更推不出它为真，至多说并非每个克里特人总说谎。另外，在历史上，这个原来认为是悖论而实际上不是悖论这一点也早已为人们所觉察，并设法修改它。最先是Eubulides（公元前4世纪）把上述命题改述为：现在我说的是一句假话。这

就是我们下文所要分析讨论的"强化了的扯谎者悖论"，对此暂且不谈。又有罗素指出，如果假定了在此克里特人说这句话之前的每个克里特人所说的每句话皆为假话这样一个前提，则上述原始命题便构成悖论。因由原始命题的真导至它为假已如上述，现假定原始命题为假，则至少有一个克里特人说过一句真话，但因为有了如上这样一个前提，这就有且仅有这一原始命题是真话。故又由它的假导至了它为真。这就构成了现代意义下的一个悖论。但是这样一个前提太强，无非是在人为地制造悖论罢了。

再举一个例子，一个人说："上帝是全能的，全能就是胜过一切"。试问此话真、假如何？设其为真，则可问："上帝能否创造一个对手来击败上帝呢？"如果能，则上帝就要被上帝自己创造出来的对手击败，故上帝并非全能。如果不能，就说明上帝还有事情做不到，即并非全能。不论何说，均导至上帝全能这句话为假。但是，反过来设这句话为假，则并不导至任何矛盾，全能的上帝本来就不存在。

以上二例均不构成悖论，却指明了一个逻辑推理，即当否定者自身被包括在被否定的对象中时，则否定者必然走向它的反面。上帝全能原要否定一切，因为上帝自己也置身于这一切之中，结果必然否定上帝自己。前例中那个克里特人也是一样，由于他自身就是克里特人，必然导至他自身说谎。

后来，人们顺着 Epimenides 的原始命题，终于构造了等价于上述 Epimenides 命题的强化了的扯谎者悖论，即"永恒性说谎者悖论"，陈述如下：

（在本页本行里所写的那句话是谎话）

由于上一行里除了这句话本身之外别无任何其他的话，因此，这是一句话中套话的话，而被套之话就是套它之话自身，于是若设该话为真，则要承认该话之结论，从而导至该话为谎话。若设该话为谎话，则应肯定该话结论的反面为真，因之推出该话为真。

这就真正构成了现代意义下的一个悖论。问题出在什么地

方？这是由于语言层次的混乱，被论断是真是假的话与去论断它的话混而为一。如果在下一行写上一句话：

（A）现在在下雨。

然后再在下一行写一句话：

（B）前一行里写的那句话是谎话。

这就不能构成悖论。话（B）是真是假，要看（A）的真假，而（A）的真假决定于现在是否下雨。在这里，话（A）是被论断的话，而话（B）是对（A）去作论断的话。故“强化了的撒谎者悖论”的症结就在于作论断的话与被论断的话混而为一。所以，这个悖论的排除在于语言的分层，这正是近代语义学产生发展的原因，也正是语义学所研究的重要内容。

最近美国数理逻辑学家Smullyan写了一本书，其书名为“这本书的书名叫什么？”现把这本书放在桌子上，并有甲、乙、丙三个人围着它，因为甲不识字而指着该书问：“这本书的书名叫什么？”而乙又接着指着这本书说：“这本书的书名叫什么？”这时丙在考虑，刚才乙讲这句话算是回答甲的问题呢？还是在重复甲的提问呢？而且乙可以叫别人永远猜不着。因为这既是问也是答。但是，如果乙诚心想回答甲的问题而又愿意把话说得清楚些：“这本书的书名就叫做《这本书的书名叫什么？》”。这就不致引起丙的烦恼了。在这里，同样存在着语义学的问题。

§2　悖论举例和数学三次危机

悖论的出现，本来并没有引起数学家和逻辑学家的重视，似乎古昔相传的悖论只是人为地制造出来的话和事，并不值得介意。直到十九世纪九十年代，悖论在数学的基础学科集合论中出现了，这样才开始引起数学家的注意，例如，下述Burali-Forti悖论最早是集论的创始者康托在1895年发现的，但他没有公开，继之由Burali－Forti在1897年发现。如所知，在超限数论中有

如下的一些定理.

定理1 任何一个良序集A不能与A的任何截段A_a相似.

定理2 凡由序数所组成的集，按其大小为序排列时，必为一良序集.

定理3 一切小于序数a的序数所组成的良序集W_a的序数$\overline{W}_a$就是序数a，即$\overline{W}_a = a$.

现将一切序数汇集在一起组成一集，记为Γ，这由康托用以造集的概括原则是可以办到的. 于是，由定理2，Γ可编成良序集，故Γ有一序数γ，即$\overline{\Gamma} = \gamma$. 既然$\Gamma$是一切序数所组成的良序集，故$\gamma \in \Gamma$，于是$\Gamma_\gamma$便是$\Gamma$之一元素$\gamma$截$\Gamma$所得的一个截段，由定理3，$\overline{\Gamma_\gamma} = \gamma$，因此，由$\overline{\Gamma} = \gamma$和$\overline{\Gamma_\gamma} = \gamma$而知$\overline{\Gamma} = \overline{\Gamma_\gamma}$，这表明良序集$\Gamma$的序数与它的一个截段$\Gamma_\gamma$的序数相同，即$\Gamma$与$\Gamma_\gamma$相似，矛盾于定理1，这就是Burali-Forti悖论.

继此之后，过了两年，康托在1899年又发现了一个悖论，人们称之为康托悖论，现陈述如下：

如所知，在古典集论中有所谓

康托定理 任何集合M的基数$\overline{\overline{M}}$小于幂集$P(M)$的基数$\overline{\overline{P(M)}}$. 其中所谓幂集，就是任给一集$M$，由$M$的一切子集所组成的集合称为$M$的幂集，并记为$P(M)$.

这样一来，根据集合论的概括原则，可有一切集合所组成的集合u. 由康托定理知$\overline{\overline{u}} < \overline{\overline{P(u)}}$，另一方面，又可证$P(u)$为$u$的一个子集. 事实上，设$x \in P(u)$，则$x$为$u$的一个子集，故$x$为一集合，故$x \in u$. 因之$P(u) \subseteq u$，即$P(u)$为$u$的子集，从而$P(u)$的基数小于等于$u$的基数，即$\overline{\overline{P(u)}} \leqslant \overline{\overline{u}}$，矛盾. 这就是康托悖论.

直到1900年为止，虽然上述集合论中的悖论出现了，但是，一则由于康托对此不加公开，二则也并不引起知情的数学家的不安，认为这仅仅是牵涉到集合论中一些较为专门的技术性问题，只要把一些定理的证明作些调整和修改，便可解决问题，从而在当时竟形成了一个充满着安全感的局面，亦即：

集合论的概念是逻辑概念，而且一般认为集合是属于逻辑的，逻辑的理论似乎应该是没有矛盾的。因此，归约到了集合论，看来快要达到目的了。的确，1900年，在巴黎召开的国际数学会议上，法国大数学家庞卡莱宣称："数学的严格性，看来今天才可说是实现了"。事实上，当时的数学家都喜气洋洋，非常乐观。

以前，对于非欧几何的不矛盾性，欧氏几何的不矛盾性，实数论的不矛盾性等等，人们虽然不能马上作出证明，但大家都相信不会导致矛盾，事实上也从未遇到出现矛盾的麻烦。

现在已把这些理论的不矛盾性直接间接地归约到集合论的不矛盾性，以致人们更加相信集合论绝不会出现矛盾。

但是，安全的想象为时不长，事隔不到两年，罗素悖论出现了，这可惊动了整个西方哲学界、逻辑学界和数学界。因为人们对罗素悖论稍加分析，就发现了只要用逻辑术语来替代集合论术语，罗素悖论就要直接牵涉到逻辑理论本身，从而直接冲击了数学和逻辑学这两门一向认为严谨的学科，这就不能不使悖论问题成为数学界和逻辑学界认真研究的课题，现将罗素悖论陈述如下：

集合可分为两种：一种是本身分子集，例如，一切概念所组成的集，由于它本身也是一个概念，所以必为该集自身的一个元素。又如一切集合所组成的集合也是一个本身分子集。另一种是非本身分子集，例如，自然数集合N决不是某个自然数n。这样，任给一集M，它不是本身分子集就是非本身分子集，不应有其他例外，现考虑一切非本身分子集的集Σ，试问Σ是哪一种集合？若设Σ为本身分子集，则Σ为自身的一个元素，而Σ之每一元皆为非本身分子集，故Σ亦应是一个非本身分子集，再设Σ为非本身分子集，而一切非本身分子集皆在Σ之中，故Σ亦应在其中，因之Σ又是一个本身分子集，不论哪种说法都说不通。这就是罗素悖论。

罗素悖论的出现，不仅对庞卡莱关于完全的严格性已经达到的说法是一个否定，而且直接动摇了他企图通过集论来为分析奠基的信心。又如德国数学家Dedekind当时正在为分析数学和数论奠基而著述《什么是数，其意义如何？》一文，逻辑学家兼数学家Frege的巨著《关于算术概念》也接近完成，由于他们的理论都涉及集合概念而延搁了上述著作的出版。

罗素悖论作为被包含在古典集合论里的一个悖论，不仅很快发现它可化归为最基本的逻辑概念的形式，而且进一步发现能用日常语言来表述它的基本原则，罗素自己就在1919年把它改写为著名的理发师悖论。现陈述如下：

李家村上所有有刮胡子习惯的人可分为两类，一类是自己给自己刮胡子的，另一类则是自己不给自己刮胡子的，李家村上有一个有刮胡子习惯的理发师自己约定："给而且只给村子里自己不给自己刮胡子的人刮胡子"。现在要问这个理发师是属于哪一类人？如果说他是属于自己给自己刮胡子的一类，则按他自己的约定，他不应该给他自己刮胡子，因之是一个自己不给自己刮胡子的人。再设他是属于自己不给自己刮胡子的一类，则按他自己的约定，他必须给他自己刮胡子，因之他又是一个自己给自己刮胡子的人了。种种说法都不通，这就是所谓理发师的悖论。

当然，罗素悖论及其变形是大家所熟悉的，由于它的重要性不能不予叙述。顺便指出，Zermelo也曾同时独立地发现了这个悖论，所以也有Russell－Zermelo悖论之称。此后，西方数学界和逻辑学界就相继出现了各种其他形式的悖论。再举一、二例。

Richard悖论在1905年发现，今选定一种语言，例如英语，则任一英语的语句，它总是由26个拉丁字母中的一些字母（可重复出现）、单字与单字之间的空档和逗点所组成的一个符号序列，而且是有穷长的，否则这句话便是一句永远说不完的话而不成其为语句了。总之，我们把任一英语的语句或一段话理解为由

28个符号中任选并可重复出现的有穷序列。但是，由一切这样的元——这种元，能用一有限或可数的符号系中的有限多个府号来表示——所组成的集是可数的，这里，在有限符号系的情况下并容许任意长的符号复合。

因此，所有用有限多个字母拼写成的任意长的字（即有意义或无意义的字母的复合）所组成的集是可数的。同样，由一切书籍，一切交响乐章以及诸如此类的东西组成的集也是如此。因此，一切英语的语句集的势是可数的。

今考虑所有能用有限句英语的语句陈述的十进位小数所组成的集E，当然$\bar{E}\leqslant\aleph_0$，如此E的一切元便可排成可数序列

$$E: \theta_1, \theta_2, \theta_3, \cdots, \theta_n, \cdots,$$

现定义一个十进位小数N如下：“如果E中第n个小数的第n位小数的数字是p，则规定N的第n位小数的数字是$p+1$，而当$p=9$时，则为0”〔英译为：*Let N be a number definied as follows. If the nth figure in the nth decimal is p, let the figure in N be p+1 (or 0, if p=9)*〕。因此，N是一个能用有限句英语的语句所陈述的十进位小数，故$N\in E$。但又由N的定义可知N与E中之每一个十进位小数皆有一有穷差位。故按有穷差位判别法可知，N与E中之一切元皆相异，故$N\notin E$。矛盾。此即Richard悖论。

基于所谓有穷可定义概念（Dixon, 1906）可以构造出与Richard悖论同类型的一系列悖论，举二例如下：

（一）试考虑一切能用有限句英语语句陈述的超穷序数的集L，并简称L中之任一元为可定义的超穷序数，完全类似的可知$\bar{\bar{L}}=\aleph_0$，但超穷序数却有不可数多，那么不妨称那些不属于L的超穷序数为不可定义的超穷序数。这些不可定义的超穷序数中必有一个最小的，设为r，于是$r\notin L$。另方面，我们就用“*the least indefinable ordinal*”〔最小的不可定义序数〕来陈述r，则应有$r\in L$。矛盾。此为一例。

（二）由于每一个自然数都能用一个英文词汇或短语去描述它，而每个英文词汇或短语总是用有限多个英文字母组成的，其中凡重复几次出现的字母就同时算几个字母，例如，100是由h、u、n、d、r、e、d七个字母描述的，同时也可用"直次于101之前的自然数"或"大于99的最小自然数"等等英文的短语去描述，那么，在各种各样的描述中总有一种所用的字母最少，特称这种描述为最简描述。今考虑："一切至多只用100个英文字母就能描述的自然数的集S"可证S为有限集合。事实上，对于每一个英文字母，可有26种选择，或者不选，计有27种选择。因之，充其量只有27^{100}种可能的描述，就算其中每一种描述都是上述意义下的最简描述。那么，由这27^{100}种描述也只能描述有限多个自然数。因之，S之元素的个数必为有限。令N为全体自然数的集，故有$M=N\setminus S\neq\emptyset$，既然$M$为非空自然数集，$M$中必有最小自然数$m$（良序集之任一子集必有首元），因此，一方面由于$m\in M$而知〔$m$是一个仅用少于或等于100个英文字母所不能描述的自然数〕。但是，我们可用下面一句话来描述m；用至多100个字母所不能描述的自然数中最小的那个自然数。现在把这句话译成英文；〔*the least positive integer which cannot be described in at most hundred letters*〕，只要数一下即知只用了67个字母〔m是一个用少于100个字母所能描述的自然数〕，故$m\in S$，于是$m\notin M$，矛盾。总之，这是用少于100个英文字母去描述了一个少于或等于100个字母所不能描述的自然数。这是与Richard悖论同类型之悖论的又一例。

我们再讲一个Grelling悖论，它在1908年被发现。在我们日常语言中，一个形容词可按其所描述的性质是否符合于该形容词本身而分为两类；凡是符合的称为自状的，凡是不符合的称为非自状的。例如，形容词"中文的"是自状的，因为不仅可以指着任何中文的书、语句或词等等说："这是中文的"，同时也可指着"中文的"三个字说："这是中文的"，即可用它自身来形

容它自身，或者说形容词所形容的性质与它自身相符合，所以它是自状的。反之，“英文的”却是非自状的，我们能指着“book”说：“这是英文的”，却不能指着“英文的”三个字说：“这是英文的”，因为它明明是中文字，亦即“英文的”这个形容词所描述的性质与它自身不相符合。又如“字”是自状的，因为它本身也是一个字。但“圆的”是非自状的，因为它本身不是圆的，而是方块的。如此，则所有的形容词被划分为自状的与非自状的两大类。但“非自状的”也是个形容词，试问它属于那一类。若设形容词“非自状的”是非自状的，则正好与它自身相符合，那么按定义它应该是自状的。再设“非自状的”是自状的，则正好与它自身不相符合，则它又应该是非自状的。不论哪种说法都说不通，这就是Grelling悖论。

还有许多其他形式的悖论，我们就不一一列举了。但是，不能不把与悖论问题相关的并称为数学史上三次危机的大事作一概要的陈述。

公元前五世纪，一个希腊人，Pythagoras学派的希帕索斯，发现了等腰直角三角形的直角边与斜边不可通约，从而导致了数学的第一次危机。事情是这样的，当时人们还处在刚刚从自然数概念脱胎而形成有理数概念的早期阶段，对于无理数的概念是一无所知。因此，当时人们的普遍见解是确信一切量都可以用有理数来表示。亦就是说，在任何精确度的范围内的任何量，总能表示为有理数，这在当时已成为希腊人的一种普遍信仰。这就是Pythagoras学派形成其观点和信条的前提。在毕氏学派看来，不仅深信数的和谐与数是万物的本源，而宇宙间的一切现象都能归结为整数或整数比已成为他们的信条。如此，上述希帕索斯的这一发现，就成为荒谬和违反常识的事，不仅严重触犯了毕氏学派的信条，同时冲击了当时希腊人的普遍见解，不能不使人们感到惊奇不安，同时也可看作是一种悖论的出现，所以这一事件堪称为数学史上的第一次危机。相传毕氏学派就因这一发现而把

希帕索斯投入海中处死，因为他在宇宙间搞出了一个直接否定毕氏信条的怪物。但是希帕索斯的伟大发现却是淹不死的，它以顽强的生命力被广为流传，迫使人们去认识和理解自然数及其比（有理数）不能包括一切几何量，迫使毕氏学派承认这一悖论和提出单子概念去解决这一悖论。单子概念是一种如此之小的度量单位，以致本身是不可度量却又要保持为一种单位。这或许是企图通过无限来解决问题的最早努力。但是，毕氏学派的努力却又引起了芝诺认为一个单子或者是0或者不是0，如果是0，则无穷多个单子相加也产生不了长度，如果不是0，则由无穷多个单子组成的有限长线段应该是无限长的，不论何说都矛盾。所以，连同著名的芝诺悖论在内也都列为数学第一次危机的组成部分。另一方面，又由于上述可称为希帕索斯悖论的出现，进一步促使人们从依靠直觉、经验而转向依靠证明，导至了公理几何学与逻辑学的诞生。

数学史上把十八世纪微积分诞生以来在数学界出现的混乱局面称为数学的第二次危机。如所知，在十七世纪和整个十八世纪由于微积分理论的产生及其在各个领域里的广泛应用，使得微积分理论得到了飞速的发展。但在另一方面，整个微积分理论却建立在含糊不清的无穷小概念上，从而没有一个牢固的基础，遭到了来自各个方面的非难和攻击。其中以Berkeley大主教对微积分的攻击最为激烈。Berkeley批判了牛顿的许多论点，例如，在《求曲边形的面积》一文中，牛顿说他避免了无穷小，他给x以增量0，展开$(x+0)^n$，减去x^n，再除以0，求出x^n的增量与x的增量的比，然后扔掉0的项，从而得到x^n的流数。Berkeley说牛顿首先给x一个增量，然后让它是0，这违背了背反律，至于导数被当作y与x消失了的增量之比，即dy与dx之比，Berkeley说它们既不是有限量也不是无穷小量，但又不是无。这些变化率只不过是消失了的量的鬼魂。Berkeley之激烈攻击微积分主要是出于他极端恐惧于当时自然科学的发展所造成的对宗教信仰的日益增长的威

胁。但也正由于当时的微积分理论没有一个牢固的基础，致使来自各方面的非难和攻击似乎言之有理。历史上，曾称Berkeley如上之论述为Berkeley悖论，而且迫使数学家不能不认真对待这个悖论，借以解除数学的第二次危机，柯西详细而有系统地发展极限论，Dedekind在实数论的基础上证明极限论的基本定理，还有康托与外尔斯特拉斯都加入了为微积分理论寻找牢固基础的工作，发展了极限理论。

普遍认为，由于严格的微积分理论的建立，上述数学史上的两次危机已经解决。但在事实上，建立严格的分析理论是以实数理论为基础的，而要建立严格的实数理论又必须以集合论为基础，而集合论的诞生与发展，却又偏偏出现了一系列的悖论，由此而构成了更大的危机。在今天，人们恰当地把集合论悖论的出现及其所引起的争论局面称之为数学的第三次危机。因此，在一定程度上讲，数学第三次危机乃是前两次危机的发展和深化，因为集合论的悖论所涉及的问题更加深刻，涉及的范围更为广阔。本篇的内容实际上就是从某一角度全面讨论数学的第三次危机，因为我们在这里始终贯穿了对这次危机的产生和寻找解决方案等情况的讨论和描述。

§3 策莫洛对悖论的解决方案

随着悖论的出现和研究，推动了人们从逻辑和哲学的角度深入研究数学基础中的问题，并取得了积极的成果。既然集合论出现了矛盾，人们当然可以把数学建基于别的理论之上，而把集合论彻底抛弃，但是，经过探索，发觉别的理论更不好弄，更难于运用，不及集合论方便有力，所以大家都致力于集合论的改造了。改造的方案主要有二：一是罗素的类型论，二是策莫洛(Zermelo, E. 1871-1953)的公理集合论。

应当指出，罗素对于悖论的研究很有贡献，这是许多数学家

和逻辑学家所公认的，虽然他的理论在实践上造成很大的困难，但是现有的一些解决悖论的方法，无不渊源于罗素早年提出的见解。因为罗素是从本质上而不是从个别技术性细节上来分析悖论的，亦即罗素把康托集合论所导致的悖论剥去了一切数学上技术性的枝节，从而揭示了这样一个惊人的事实，即我们的逻辑直觉（诸如关于真理、概念、存在、集合的直觉）是自我矛盾的。罗素认为一个集合可以用两种方法予以定义，我们可以枚举它的元素，……，也可以指明它的性质，……，那种枚举式的定义称为外延性的定义，而那种指明性质的定义方法称为内函式的定义。基于这样一种考虑问题的准则，寻找出路的办法有两条道路可走，这就是罗素当年所指出的几个可能方向中的，罗素称为**量性限制理论**和**曲折理论**这样两个方向，更确切地可称之为**外延理论**和**内函理论**。对于罗素指出的曲折理论，后来在Quine的工作中得到了阐发，至于量性限制理论，其最主要的特性就是对于全集或无限制的关于某种现象的概念的存在性加以限制，后来由策莫洛和其他人所发展起来的公理化集合论在涉及集合时就可以看成是对这一思想的阐发。现在，我们就来分析讨论策莫洛等人在这方面的工作及其思想方法。

自从罗素悖论出现以后，策莫洛想借助于他所说的划分公理（或称分出公理）来排除它，其法是否有效，尚要看看他的具体做法再作定论。

划分公理 设L为任一集合，$R(\theta)$是与L的变元有关的一句话，则L中一切能使$R(\theta)$成真话的元素可组成一集合$L_{R(\theta)}$。

显然，$L_{R(\theta)}\subseteq L$，即$L_{R(\theta)}$为L的一个子集。

根据划分公理，可证下述定理为真。

定理 任意一集L必有一子集不是L的元素。

证明 设$\theta\notin\theta$，其余符号意义如公理中所述。现在只要证明$L_{R(\theta)}\notin L$即可。因为这就找到了L的一个子集$L_{R(\theta)}$不是L的元素。现用反证法，假设$L_{R(\theta)}\in L$，则作为L之一元的$L_{R(\theta)}$，必或

使$\theta\notin\theta$为真，或使$\theta\notin\theta$为假。因此，$L_{R(\theta)}\in L_{R(\theta)}$与$L_{R(\theta)}\notin L_{R(\theta)}$两个关系中有且仅有一个成立。但是，若$L_{R(\theta)}\in L_{R(\theta)}$，则因$L_{R(\theta)}$中的每一元皆使$\theta\notin\theta$为真，于是，作为$L_{R(\theta)}$之一元的$L_{R(\theta)}$应使$\theta\notin\theta$为真，故推得$L_{R(\theta)}\notin L_{R(\theta)}$。再设$L_{R(\theta)}\notin L_{R(\theta)}$，而这正表示作为$L$之一元素的$L_{R(\theta)}$使$\theta\notin\theta$成真话，从而必在$L_{R(\theta)}$之中，故又推得$L_{R(\theta)}\in L_{R(\theta)}$。条条路都行不通。说明原先的假设$L_{R(\theta)}\in L$一事不得成立。故$L_{R(\theta)}\subseteq L$但$L_{R(\theta)}\notin L$。证毕。

有了这条定理，即可证明罗素悖论陈述中的那个"一切非本身分子集的集Σ"不是一个集合。否则，设Σ为一集，则由定理可知，Σ必有一子集Γ不是Σ的元素，即有

$$\Gamma\subseteq\Sigma \qquad 且\Gamma\notin\Sigma \quad (\bigstar)$$

作为集合Σ的子集Γ必为一集，故可问Γ是本身分子集还是非本身分子集，设Γ为本身分子集，则有$\Gamma\in\Gamma$，但$\Gamma\subseteq\Sigma$，故$\Gamma\in\Sigma$，这与（★）矛盾。再设Γ为非本身分子集，则因Σ为一切非本身分子集的集，故$\Gamma\in\Sigma$，又矛盾于（★）。哪种说法都不通，说明Σ为集的假定不得成立。故Σ不是集合。既然如此，就不能问Σ是本身分子集还是非本身分子集。

这样一来，从表面上看，似乎只要承认划分公理，罗素悖论就能排除，其实并非这样简单。首先要问，划分公理是和哪些原始概念、推理原则和基本原理结合起来进行推理的。否则，仅由一条划分公理，不要说进行推理，连它自身的陈述也不可能。所以，只要承认划分公理，便能排除罗素悖论这样一种孤立的说法，不只是粗糙的，可以说是错误的。在此，需要一整套的原始概念和推理原则，即要有它的公理系统。对此先让我们明确此点而姑且暂缓考虑。

现在，还是先让我们认真考虑一下，问题究竟出在什么地方？如所知，概括原则是康托建立古典集合论的重要思想方法之一，但是，稍加分析，即可看出，上述那个划分公理不过是概括

原则的一个简单推论，只要把集合L且$R(\theta)$作为一个性质P，则$L_{R(\theta)}$是一集，在概括原则之下是理所当然的。如此，在康托系统中，一方面由概括原则推知上述那个Σ（一切非本身分子集的集）是一个集合，另一方面，由概括原则而推出划分公理，又如由上划分公理而推出Σ不是集合。由此可见，概括原则本身就是自相矛盾的。

再举一例，一方面由概括原则推出划分公理，又由划分公理推知任一集合必有一子集不是它的元素。从而有下面的结论：

（$*_1$）一切集合所组成的集合E不是集合。否则，设E为一集，则E有一子集Γ不为E的元素，因E被假定为集，故E的子集Γ必为集合，而$\Gamma \notin E$表示E并非一切集的集，至少把Γ漏掉了，从而与E的定义相矛盾。证毕。

另一方面，由概括原则：

$$G=\{g \setminus p(g)\} \qquad 和 \forall g(g \in G \longleftrightarrow p(g))，$$

只要令“集合”为一性质，即有

$$E=\{s \setminus s为一集\}和\forall s(s \in E \longleftrightarrow s为一集)。$$

因此，又有结论：

（$*_2$）一切集合所组成的集合E是一个集合。

如此，（$*_1$）和（$*_2$）均由概括原则推出，但它们是互相矛盾的。可见，矛盾的交点已集中到概括原则这个康托用以造集的思想方法上去了。

后来，经过仔细分析，人们归结到如下四件事不能同时成立，这就是：

（1）$x \notin x$是一个条件（含x的语句）。

（2）任给一条件$\phi(x)$决定一集，即$x \in A \longleftrightarrow \phi(x)$。

（3）集合为个体之一，因而x处均可代以A。

（4）$p \longleftrightarrow \neg p$为一矛盾。

现在我们来证明上述（1）、（2）、（3）、（4）不得同时成立。否则，假设它们同时成立，则由（1）知$x \notin x$是一个条件，

故我们可取$x \notin x$为一条件$\phi(x)$，由（2）而有$x \in A \longleftrightarrow x \notin x$，由（3）而知$A$可代入$x$处，于是$A \in A \longleftrightarrow A \notin A$，由（4）而知$A \in A \longleftrightarrow A \notin A$是一矛盾。证毕。

所以，在（1）、（2）、（3）、（4）中至少要否定一条。罗素从否定（1）出发而展开他的类型论，Zermelo－Fraenkel基于否定（2）而构造ZFC集合论公理系统，Bernays－Gödel在否定（3）的基础上形成他们的BG集合论公理系统，Бочвар以否定（4）为起点，发展他的多值逻辑。其中ZFC系统和BG系统本质上差不多，实质上均以修改概括原则为前提，即在不同程度上对概括原则的那种造集的任意性加以适当的限制，到目前为止，单从排除悖论这一点讲，类型论、ZFC系统、BG系统都能使已经发生的逻辑、数学悖论不在他们的系统中出现。至于那些语义学悖论，则又当别论，已在Tarski的工作中获得一定的处理。那么多值逻辑的展开能否排出悖论呢？在承认概括原则的前提下，多值逻辑的展开对于悖论的排除来说，将是无济于事的。当然，发展多值逻辑的价值，完全取决于能不能排除悖论这一标准。

有了如上的简要综述，即可把我们的讨论再转向策莫洛的工作。如前所述，矛盾的交点已集中到概括原则本身，但是，概括原则这一思想规定究竟在什么地方出问题，似乎又突出地表现在"一切集合汇集在一起"究竟能不能构成集合这样一个问题。致使人们很快就把注意力集中到概括原则所肯定的那种造集的任意性这一点上。因之，策莫洛首先构造公理系统，在保留概括原则中之合理因素的前提下，对造集的任意性加以适当的限制，形成了一个包括划分公理在内的集合论公理系统。在这个系统内，只承认按系统中公理所允许的限度内构造出来的集才是集合，凡是超出系统中公理所允许的限度而构造出来的集是概不承认的。特别是有如由一切集组成的集等等在这个公理系统中是不被承认为集的。在这个系统中能把康托悖论、Burali－Forti悖论以及罗

素悖论等等已经出现的逻辑、数学悖论予以排除。

这样在策莫洛于1908年建立了他的集论公理系统后，Fraenkel与Skolem在1921—1923年间给出了一个严格的解释，并对策莫洛公理系统作了改进，形成了今天著名的ZF系统，加上选择公理，便是熟知的ZFC系统。

现在，我们除ZF的形式语言和一阶谓词演算的公理和推理原则略而不叙外，将ZFC系统的诸非逻辑公理陈述如下，以对这一系统的基本原则有一概略了解。最后，我们再对ZFC系统作一简短的一般性评述。ZFC系统的陈述有许多版本，在这里，我们采用Herbert B. Enderton所著《ELEMENTS OF SET THEORY》一书中的陈述形式，在该书之末所附的公理表中，包括了外延、空集、配对、并集、幂集、子集、无穷、选择、代换、正则等十条公理，我们不仅要按这些公理在该书中的出处一一笔录，而且特别重视这些公理的非形式化的自然语言的表达方式。

外延性公理 如果两个集合有完全相同的元素，则它们相等：

$$\forall A \forall B [\forall x (x \in A \Longleftrightarrow x \in B) \Longrightarrow A = B]$$

空集公理 存在一个不含任何元素的集合：

$$\exists \varnothing \forall x (x \notin \varnothing)$$

对偶公理 对任何集合u和v，存在一个集合恰以u和v为其元素：

$$\forall u \forall v \exists B \forall x \ (x \in B \Longleftrightarrow x = u \ \text{or} \ x = v)$$

并集公理 （初级形式） 对任何两个集合a和b，存在一个集合，它的元素或者属于a或者属于b（或者属于两者）的集合：

$$\forall a \forall b \exists B \forall x \ (x \in B \Longleftrightarrow x \in a \ \text{or} \ x \in b).$$

幂集公理 对任何集合a存在一个集合，它的元素恰好是a的所有的子集：

$$\forall a \exists B \forall x \ (x \in B \Longleftrightarrow x \subseteq a)$$

只要愿意，这里的“$x\subseteq a$”可以改写为$\in$的定义形式：

$$\forall t\ (t\in x\Longleftrightarrow t\in a).$$

以后，我们要扩大这组公理，包括子集公理、无限公理、选择公理、代换公理和正则公理。并集公理还要用更强的形式重新叙述（并非全部这些公理都是真正必需的，将会发现其中某些是多余的）。

子集公理 对每一个不包含B的公式——，下式

$\forall t_1\cdots\forall t_k\forall c\exists B\forall x\ (x\in B\Longleftrightarrow x\in c\&——)$ 是公理。

若用语言叙述，这个公理断言（对任何t_1，…，t_k和C）存在这样的一个集合B，其元素正好是C中所有使____成立的那些集合x。于是，自然得出B是C的子集的结论（子集公理之名由此而来）。集合B被（t_1，…，t_k和C）唯一确定，并可用抽象记号的变形

$$B=\{x\in C\mid ____\}$$

给它命名。

我们需要改进并集公理（初级形式）的说法，以便知道存在一个集合，它正好包含着A的元素的元素全体。

并集公理 对任意集合A，存在一个集合B，它的元素正好是A的元素的元素全体：

$$\forall x[x\in B\Longleftrightarrow(\exists b\in A)x\in b].$$

定义 对任何集合a，它的后继a^+定义为

$$a^+=a\cup\{a\},$$

当且仅当$\varnothing\in A$并且在后继下封闭，即

$$(\forall a\in A)a^+\in A$$

时，说集合A是**归纳集合**。

虽然我们还没给出无限的形式定义，我们可以正式看到，任何一个归纳集将是无限的。

到现在我们还没有提供无限集的存在性公理。虽然我们能确立其存在的无限多个不同的集合。但是，我们还没有能证明有无

限多元素的集合。因此，我们还不能证明任何归纳集合的存在。为此，我们引进

无限公理 存在归纳集合：

$(\exists A)\ [\varnothing \in A \& (\forall a \in A) a^{+} \in A]$。

借助于这个公理，我们能定义自然数。

选择公理有许多种等价的形式，今取其一种并用自然语言陈述之。

选择公理 令$\mathscr{A}$是这样的一种集合：(a)$\mathscr{A}$的每个元素是非空集合，并且(b)$\mathscr{A}$的任何两个不同的元素不相交。那末存在集合C恰好包含$\mathscr{A}$的每个元素中的一个元素（即对每个$B \in \mathscr{A}$，$C \cap B$是某个x的单元集$\{x\}$）。

代换公理 对任何不包含字母B的公式$\varphi(x,y)$，下述公式是公理：

$$\forall A[(\forall x \in A)\forall y_1 \forall y_2(\varphi(x,y_1)\ \&\ \varphi(x,y_2) \Longrightarrow y_1 = y_2) \Longrightarrow \exists B \forall y(y \in B \Longleftrightarrow (\exists x \in A)\varphi(x,y))]。$$

现把代换公理译成自然语言：把公式$\varphi(x,y)$读作“x提名y”则公理的前提说：“A的每一个元素至多提名一个对象”。而结论为所有被A之元提名的那些对象组成一集B。名称代换反映集合A中的每一个x由它的被提名者（如果有的话）代替而产生集合B这样一种思想。

正则公理 每个非空集合A有元素m，且$m \cap A = \varnothing$：

$$(\forall A \neq \varnothing)(\exists m \in A) m \cap A = \varnothing。$$

如上这些公理便构成ZFC集合论公理系统，从这些公理内容可看出，Zermelo－Fraenkel构造ZFC系统，其真实目的乃在于为分析学奠定严格的基础，因为如前所述，微积分的基础已通过柯西到Dedekind归约到了实数论，而实数论的不矛盾性又归约于集合论的不矛盾性。那么ZFC系统则以如下路线来为微积分奠基，这就是由无穷公理来保证自然数集的合法性，再由幂集公理导至实数集的合法化，然后再由子集公理来保证实数集中满足性质P

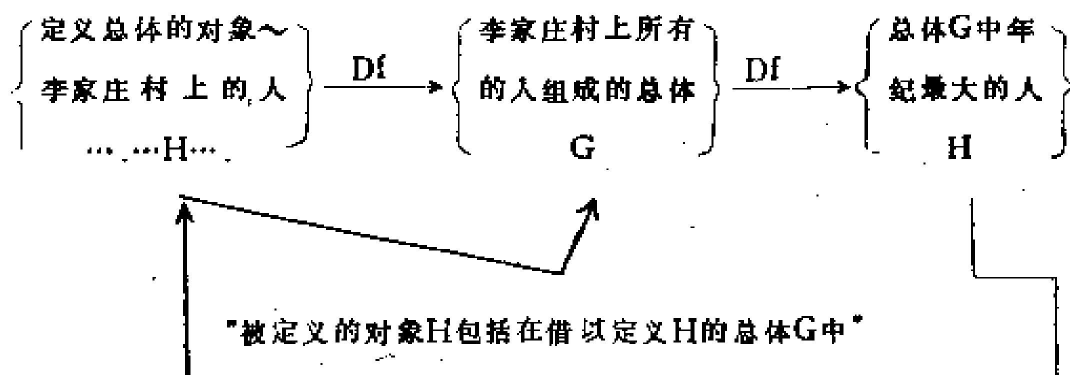
的元所组成的子集的合法性，这样一来，只要ZFC系统无矛盾，严格的微积分理论就能在ZFC公理集合论上建立起来了。但是，问题正在于ZFC系统本身的无矛盾性至今没有解决。所以至今不能保证在这个系统中今后不会出现悖论，虽然在ZFC系统中能够排除已经出现的那些集合论悖论，并且ZFC系统一直展开到今天，尚未出现过其他矛盾。但是，庞卡莱指出："我们设置栅栏，把羊群围住，免受狼的侵袭，但是很可能在围栅栏时就已经有一条狼被围在其中了"。

§4 罗素对悖论的解决方案

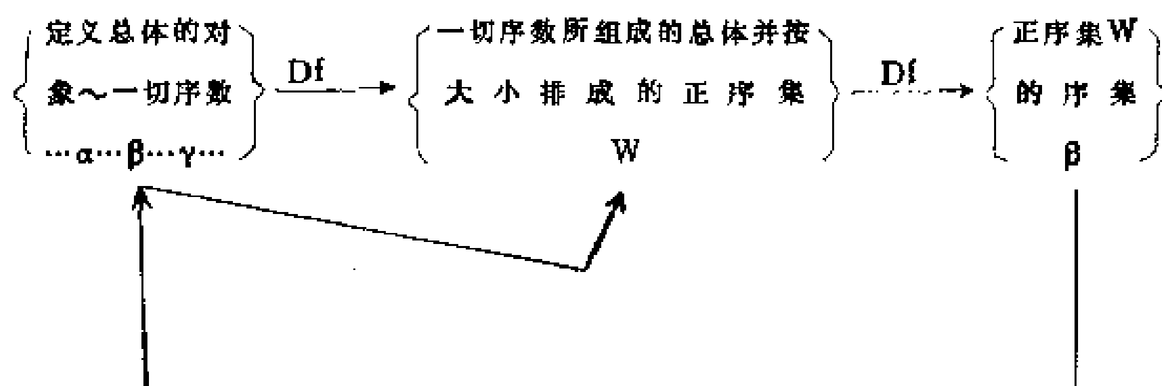
关于悖论的成因，庞卡莱曾在1905，1906，1908年多次指出，所有的悖论都与非直谓的定义有关。什么叫非直谓定义呢？就是被定义的对象被包括在借以定义它的各个对象中。说得更明确一点，就是借助于一个总体来定义一个概念。而这个概念本身又属于这一总体。举例如下：

例1 李家庄村上年纪最大的人H。

这里被定义的对象是H这个人，但当我们对H下定义时，一方面要借助"年纪最大"的概念。特别要借助于"李家庄村上所有的人所组成的总体G"这一概念。但在定义总体G的时候，首先要借助李家庄村上每一个人，其中包括H这个人在内，这就是借助总体G来定义H。而H本身又属于总体G。如图：

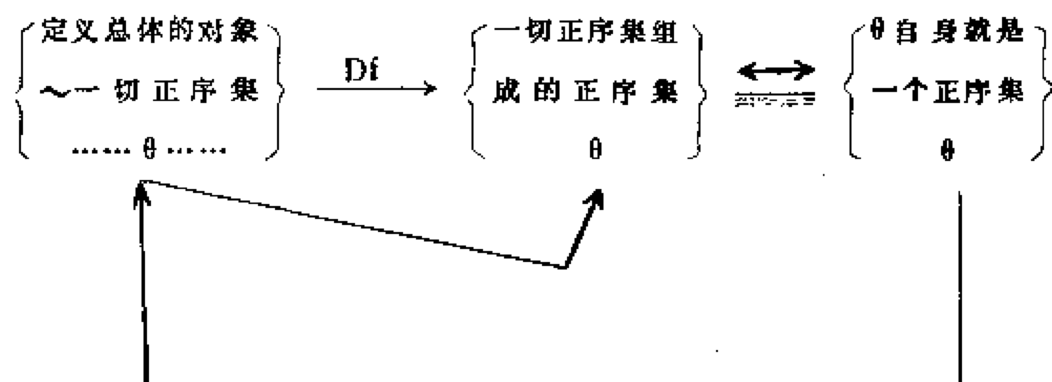


例2　一切序数所组成的正序集W的序数β。

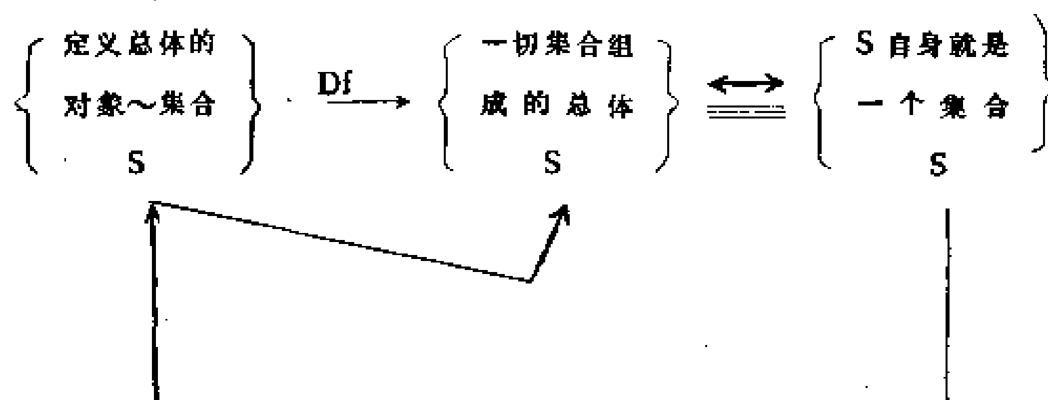


例3　一切正序集所组成的正序集θ。

每一正序集都对应于一个序数，故把一切正序集汇成总体后，再按每一正序集所对应的序数的大小为次序把这个总体排成整序集θ。



例4　一切集合所组成的集合。

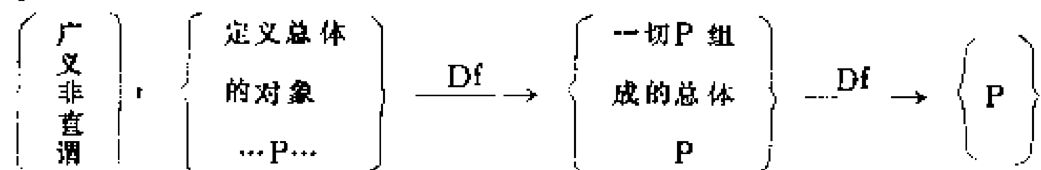


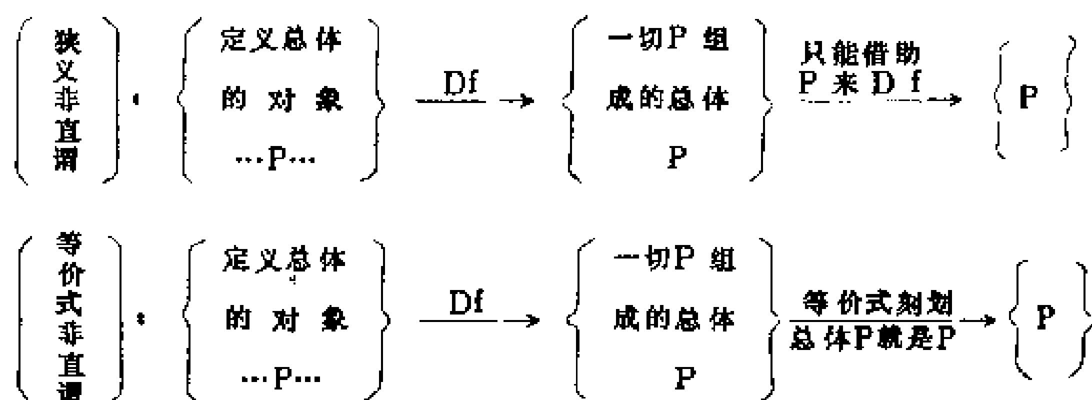
以上四例都是非直谓定义法，但如何构成非直谓定义的具体过程却各有不同，可作如下分类：

（一）广义非直谓。凡是非直谓定义中的被定义对象，可用直谓定义法重新定义者，亦即被定义对象并非只能借助于包括它的总体来加以定义者叫做广义非直谓。如例1中的H，既可借助于G来定义，也可用“李大娘的老伴”、“五小二的祖父”等等不借助于总体G来直谓地定义H。

（二）狭义非直谓。凡是非直谓定义中的被定义对象非借助于总体不可，亦即被定义的对象只能借助于这一总体才能定义。如例2中的序数β。由于序数β就是正序集W的序数β。如果丢开总体W，如何来讲W的序数呢？亦即因为β是W的β，不是别的β。那么β就只能借助于W来定义它了。让我们再回到例1，设想有一个不是李家庄村上的外地人K，并且K对李家庄村上的人和事一无所知。现在K要到李家庄村上找年纪最大的人H了解一件事，那么K到李家庄后，如何向李家庄村上的人来表达“他要找H问事”这个意思呢？在此时此刻对这个特定的人K来说，就非借助于总体G来向别人询找H这个人了。因此，例1本来是广义非直谓，但相对于特定条件下的K而言，就成为狭义的非直谓了。

（三）等价式的非直谓。凡是非直谓定义中的被定义对象仅借助于“总体本身就是什么”这样的等价式刻划来确定的，则称为等价式的非直谓。如例3中的被定义对象就是通过“θ本身就是一个正序集”这样的等价式刻划来确定的。例4中的被定义对象也是通过“一切集合的集合就是一个集合”这种自身等于自己的方式来刻划的。如果把这种等价式的刻划亦算作一种自相的定义方式的话，当然亦必须纳入非借助于总体定义的情形中去，因此，等价式非直谓是狭义非直谓的特殊情形，而狭义非直谓又是广义非直谓的特殊情形。综上所述，可将非直谓的三种情形图示如下：





罗素自己在解决悖论问题的征途上，并没有沿着他自己所指出的那两个方向前进，而把自己的工作健基于对悖论的更为一般的分析上，那就是对非直谓定义的分析，鉴于这一分析，罗素进一步明确了庞加莱关于悖论成因的想法，认为所有这些悖论，都有这样一个关键性的对象，它借助于一个整体予以定义和刻划。但这一对象却又被包含在这一整体中，在这里出现了一种循环，而正是由于这种循环导致悖论的出现。因此，构成悖论的深刻原因确与非直谓定义法有关。为之，罗素明确提出如下的原则：

恶性循环原则：没有一个整体能包含一个只能借助于这个整体来定义的元素。

基于这一原则，罗素形成和发展了他的分支类型论。

如果我们把恶性循环原则对照前面所论的关于非直谓定义法的三种情形加以分析，可看出此原则与狭义非直谓定义法直接相关，而作为狭义非直谓的一种特殊情况就是等价式非直谓定义法，从而由恶性循环原则对狭义非直谓定义法的否定也就包括了对等价式非直谓的否定。而等价式非直谓从集合的观点讲则与“一集为它自身之一元”直接相关。由此可见，罗素的恶性循环原则从集合的观念出发就直接隐含了下述一种思想规定，我们不妨称这种思想规定为：

类型混淆原则：任何一个集合决不是它自身的一个元素。

实际上，罗素首先是对于性质，要按照它所属的对象的类型加以分类，属于O类的是那些论域中的对象的名称，如a、b、c、

…。属于第一类的则是这些对象的性质，如$f(a)$，$g(b)$…中的f、g、…，属于第二类的则是那些性质的性质，如$F(f)$、$G(f)$、…中的F、G、…。属于第三类的则是性质的性质的性质等等。

作为类的划分的一个基本原则是，每一谓词（性质）都必须从属于一个确定的类。而且每一类的性质只有当其使用于直次于它的那个类的对象才是有意义的。因此，如$f(a)$、$F(g)$、…是有意义的，但$F(a)$、$f(f)$、$a(b)$、…就是无意义的。

现在，我们按照罗素上述对于类的划分及其必须遵循的原则，对于由a、b、c、…为元素而构成的集合S加以分析，为使讨论简单明了不妨把元素a、b、c、…视为论域中的对象的名称，即属于O类的，而构成集合S之后对于它的元素a、b、c、…的性质是属于第一类的。那么，按照每一类的性质只有当其使用于直次于它的那个类的对象才有意义的规定，可知只有$a\in S$、$b\in S$、…才是有意义的（相当于上述$f(a)$…）。但是，$s\in s$、$b\in c$、$c\in c$、…等都是没有意义的（相当于上述$f(f)$、$b(b)$、$c(c)$、$b(c)$…一类的情形），也是所给原则之下所不允许的。因此，集合S就决不会是S自身的一个元素，否则，若设$S=C$，则就将出现$s\in s$、$c\in c$一类违反原则而无意义的事。可见，完全符合上述类型混淆原则这样的思想规定。

罗素认为，除掉对于性质要作出类的划分之外，还要对于性质的定义方法，再把同一类中的性质作出级的划分，那些在定义方法中没有涉及到“所有的性质”的性质是第1级的，而那些在下定义时涉及到第n级的“所有的性质”的性质是第$n+1$级的，…，如此，当不具体指明所考虑的级，则凡涉及到“所有的性质”的表达式是无意义的。

这样，所有的性质（谓词）首先按对象的性质、性质的性质…加以分类。再按类中性质的定义方法加以分级。因之，每一性质都归属于一定的类和级，由于级是在类内划分的，因此，罗素

的这一理论就叫做分支类型论。

在这里，首先要注意级的划分是对定义方法而说的，而类的划分则不是，所以两个划分在含义上完全不同。其次，就逻辑语言中的性质（谓词）而言，也不要单纯从日常语言的角度去作字面上的理解。这样我们就可通过前述关于非直谓定义的各个举例来把上面“类中分级”那一段文字讲明白。让我们回到例1，现把李家庄村上每一个人看作是某一类中的第n级的一个性质，那么，按例1中那种非直谓定义法对H下定义时，曾借助了李家庄村上所有性质（人）所组成的总体G，因此，H就是一个涉及到第n级的“所有的性质”的性质了。按规定H就是第$n+1$级的性质了。如此，按照恶性循环原则，凡是只能借助于第n级的“所有性质”来定义的那个第$n+1$级的性质就决不允许包含在第n级的性质中。试看前述例2，不妨把所有的序数算作是第1级的性质，则一切序数所组成的正序集W的序数β这个性质就是第2级的了。而且β这个第2级的性质是一个只能借助第1级的“所有的性质”来定义的性质，因此，按照恶性循环原则，β一定不能被包括在W之中。

现在，让我们再回到前述三种类型的非直谓定义方法中去讨论罗素的恶性循环原则与类型混淆原则。首先，广义非直谓定义法不牵涉到这两个原则中的任何一个，但是狭义非直谓定义法却与恶性循环原则直接冲突，由于等价式非直谓是狭义直谓的特殊情况，因此，当承认恶性循环原则而否定狭义非直谓定义法时，等价式非直谓定义法就自然排除了。另一方面，构成等价式非直谓的关键之处在于由整体到被定义对象的等价式刻划，而等价式刻划却冲突且仅冲突于类型混淆原则。这样一来，如果只承认类型混淆原则而不承认恶性循环原则的话，则在狭义非直谓定义法中就只排除等价式非直谓的特殊情形，而可保留那些不涉及类型混淆和等价式刻划的一般意义下的狭义非直谓定义法。

实际上，那些不涉及类型混淆的非直谓定义法不仅在数学上，而且在日常生活中也常常是不可缺少的。从数学上看，外耳

曾经指出，把上确界定义为上界中最小的一个，这就是只能借助于“上界的总体”来定义上确界的非直谓定义法，但此处虽然包含了循环，却是无害的，它不仅没有引向悖论，而上确界的概念在数学中是十分有用而重要的。在日常生活中，可以前述例1中那个外地人K为例，当K到李家庄村上去找年纪最大的人H时，由于K对李家庄村上的人和事一无所知，因此，对K来讲，就非要借助于狭义非直谓定义法来表达他要找H的想法，这说明K在此时此刻就不能同时为恶性循环原则的笃信者。特别是在数学领域中，如果坚持贯彻恶性循环原则，势必要将许多数学概念抛弃，许多数学定理不能证明，从而虽然排除了悖论，却同时失去了许多合理的东西。再加上罗素的分支类型论的展开显得非常累赘、烦琐和复杂，以致不为数学家所接受。

后来，有罗素的学生兰姆赛（Ramsey）把分支类型论加以简化，在兰姆赛看来，悖论可分为两类，一类叫做逻辑，数学悖论，它只是由逻辑或数学系统中的概念构成，因之总能用逻辑和数学符号来表达。例如前述康托悖论、罗素悖论、布拉里福尔帝悖论等均属这一类。另一类可叫做语义学悖论，这类悖论由命名、真、假等概念构成，因此，如表明、刻划、真、假等语义项成为它不可少的组成部分。例如前述的永恒的撒谎者悖论、理发师悖论、格雷林悖论等等皆属这一类。综观前述逻辑数学悖论，无不导源于“一切集合的集是集”、“一切正序集的集是正序集”等等违反类型混淆原则的“本身分子集”，后经兰姆赛的研究，只要坚持类型混淆原则，便足以排除逻辑数学悖论。而对于那些语义学悖论，则由于它不能在逻辑数学的符号语言中表达，从而由逻辑出发去构造数学时，可以不考虑语义学悖论。因之，兰姆赛就在废除分支类型论中关于级的划分而保留类的划分的基础上建立了简单类型论。并为大多数数学家所欢迎。对于非直谓定义法，兰姆赛指出，这里的确包含了循环，但这一循环是无害的，不是恶性的，此处作为这一断言的认识论基础是这样的。性

质的整体本身已经存在着，而作为非直谓定义只是一种加以鉴别的方法，如同“房间里最高的人”一样，这是一个经验的事实。就是作为有限的存在不可能单独地对无穷多个性质中的每一个去进行命名，却可以通过所有性质的整体而对其中一些进行描述。所以，兰姆赛在数学的领域里仍然保留了狭义非直谓定义法的合理使用而又排除了逻辑、数学悖论的出现。如此，罗素的分支类型论就显得多余了。

§5 塔斯基及其语义学

塔斯基（Tarski）曾对语义学悖论作了比较彻底的研究。他指出，语义学悖论的成因，特别是导致诸如强化了的撒谎者悖论出现的原因在于自然语言的含糊性，自然语言的一个特徵在于它的普遍性，正是这种普遍性构成了所有语义学悖论，诸如永恒性撒谎者悖论等出现的主要原因。因为，一种语言包含了它自身的语义学概念，而在其中普遍的逻辑法则又同时适用的话，这就必然是不相容的。罗素为要排除语义学悖论，也曾建议把它的类型论扩大为分层语言论，塔斯基顺着罗素的思路发展和形成了理论语义学，展开了关于对象语言和元语言的讨论。塔斯基在他的语义学理论中指出，在一个语言系统的内部是不能定义关于这一语言的语义学概念的，对象语言的语义学概念必须在对象语言以外的元语言中予以表述，而元语言的语义学概念又必须在元语言以外的元元语言中予以表述。按照这样的原则确立语义学概念当可同时避免语义学悖论的出现，但在这里还有一个前提，那就是只有在元语言中具备此对象语言中所有的变量都具有更高逻辑类型的变量时，才能在元语言中构造出在方法论上正确并在实质上是充分的语义学概念的定义。相反的，如果对象语言也足够丰富，以致元语言中所有的概念和语法形式都能在对象语言中得到翻译。那么，上述一套方法将是无济于事，语义学的悖论在定义语

义学概念时仍将不可避免。但是塔斯基又指出："即使对无穷阶的语言来说，真理概念的无矛盾性和正确的应用仍然是可能的，只要把真理概念作为元语言的原始概念之一，并通过公理来确定它的性质"。

事实上，只有当塔斯基证明了人们可能把真理这个概念引入演绎理论中而不会陷入矛盾时，严格意义的语义学才开始它真正的发展。在塔斯基之后，直至今天，西方逻辑学界和数学界仍然在悖论问题上进行着热烈的讨论，而杰罗姆·马立兹说："这个问题，哲学家们和数学家们热烈讨论了好几十年，却至今没有得出圆满的答案"。亦就是说，悖论问题迄未获至圆满而彻底的解决。

§6 哥德尔的不完备性定理与悖论

如所知，著名的哥德尔（Gödel）不完备性定理是数理逻辑发展史上的重大研究成果，曾被誉为逻辑在现代所取得的最重要的进展之一，是数学与逻辑发展史中的一个里程碑。我们在这里将要讨论不完备性定理与悖论问题的关系。用哥德尔自己的话来讲，他获得并证明这个定理是直接来自对罗素悖论的分析。这就是说，不完备性定理的直观背景及其证明思想与悖论的分析有着密切的联系。由此可见，从方法论的角度来研究悖论问题，该有多么重要的意义。

哥德尔关于形式系统的不完备性定理首次发表在他的论文《论数学原理及有关系统中的不可判定命题》中，那个定理是关于不可判定命题存在性的一般结果，如果仅就算术系统而言，则该定理可简单地表述为：

定理 如果形式算术系统是ω无矛盾的，则存在着这样一个命题，该命题及其否定在该系统中都不能证明，亦即它是不完备的。

后来，又经罗塞尔（Rosser）改进为下述形式：

定理 如果形式算术系统是无矛盾的，则它是不完备的。详言之，亦即

定理 如果一个含有自然数论的形式系统S是无矛盾的，则S中存在一个逻辑公式A，使得在S中，A是不能证明的，同时$\neg A$也是不能证明的。

作为不完备性定理证明思想的一个关键之处在于映射思想的天才应用。哥德尔通过一种十分新颖的映射形式来建立他的主要结论。实际上，所说映射是数学领域内的一种极为重要的研究方法，其基本思想就是借助一一对应使得某一领域内的对象之间的某种关系得以在另一领域内的对象之间的关系得到表现（请参考第3章）。而哥德尔把算术系统（记为N）中的符号、表达式和表达式的序列都映射为数——他通过引进哥德尔数而实现了对象的数化手续。这样一来，对于数理逻辑及其他有关分支来说，在研究方法上就提供了一种数字化的工具。从而方便地把一些讨论对象（如符号，公式）转换为自然数或自然数的函数，致使我们能用自然数函数的理论来讨论有关问题。其次，哥德尔又通过递归函数的引进证明了所有元理论中关于表达式的结构性质的命题，均可在算术系统中得到表示。如此，就使元理论中的命题都映射为算术系统中的命题。致使算术系统中的一部分表达式获得了元数学意义。

哥德尔自己在阐明其证明思想时指出：“我们可以注意到一个形式系统的公式在形式上都表现为基本符号（变量、逻辑常项、括号或中断符号）的一个有限序列，而且人们容易精确地去指明基本符号的那些有限序列是有意义的公式和那些不是有意义的公式。类似地，从形式的观点看，所谓证明实际上就是公式的一个有限序列。对于元数学来说，究竟用什么东西来作为基本符号当然是没有关系的。我们不妨就用自然数来作为基本符号，如此，一个公式就是一个自然数的有限序列，而一个证明便是一个

有限的自然数序列的有限序列。据此，元数学的概念（命题）也就变成了关于自然数或他们的序列的基本概念（命题），从而即可（至少是部分地）在（对象）系统本身的符号中得到表示，特别是人们可以证明‘公式’、‘证明’、‘可证公式’等都可在对象系统中加以定义”。

如此，作为不完备性定理证明的第二步是在对象系统内构造这样一个命题G，使其元数学的意义为“G是不能证明的”（注意，这是元数学中的命题——我们记为G'）。

哥德尔指出，一旦构成了这样的命题，定理的证明就完成了，因为G正是所需要的不可判定的命题，现简要地描述如下：

前提： （α）凡是可证明的命题必然是真的（从直观上看，这是任一公理系统的必然要求）。

（β）命题的真理性在映射下保持不变（特别是这里的G和G'是同真假的）。

结论1 G是不能证明的。

证明 现用反证法

设G是可以证明的$\xrightarrow{(\alpha)}G$为真$\xrightarrow{(\beta)}G'$为真$\xrightarrow{\text{由}G'\text{的意义}}G$是不能证明的。矛盾，证毕。

结论2 $\neg G$也是不能证明的。

证明 由结论1可知，G是不能证明的$\xrightarrow{\text{由}G'\text{的意义}}G'$为真$\xrightarrow{(\beta)}G$为真$\xrightarrow{Df}\neg G$为假$\xrightarrow{(\alpha)}\neg G$是不能证明的，证毕。

由结论1和结论2可知G是不可判定的，亦即系统是不完备的。

哥德尔自己指出：“这一推理过程与里查德（Richard）悖论的相似之处是显然的，而且和强化了的撒谎者悖论也有一个很大的相似性。因为那个不可判定的命题〔$R(g, g)$〕所断言的正就是…〔$R(g, g)$〕是不可证明的”。关于里查德悖论，我们在§2中特地引进了关于它的“里查德数等价于非里查德数”的陈述形式，大家可以自行对照分析。

如所知，塔斯基证明了

定理 对于无穷阶的形式语言来说，如果相应的元理论中的可证明命题集是无矛盾的，那么就不可能在元语言中构造出一个在约定意义下是充分的关于真理的定义。

这是一个关于真理概念的可定义性的定理，值得注意的是塔斯基对本定理的证明思想与方法完全类似于上述哥德尔对不完备性定理的证明思想和方法。

哥德尔的定理和塔斯基的定理都具有深刻的数学和哲学意义。而且它们的证明思想都是直接来源于悖论的分析。可见从方法论的角度来看悖论问题的研究确有重要意义。

§7 悖论的成因与研究悖论的重要意义

我们在§1中论及悖论的定义时曾指出，任何一个悖论都相对于某一理论体系，不妨进一步指出，任一悖论还属于一定的历史范畴，亦即它还相对于认识的各个历史阶段。如希帕索斯悖论、伽利略悖论和贝克莱悖论等等均已进入历史博物馆。既然任何悖论都属于一定的历史范畴和一定的理论系统，也就不存在任何绝对意义下的悖论。既然如此，也就没有什么绝对意义下的产生悖论的终极原因，同时也不会有任何能在绝对意义下一劳永逸地免除悖论的方法，但这又并不排斥我们能在认识论的意义下去揭示产生悖论的根本原因。如所知，人的认识总具有历史的局限性和相对性，因之，产生悖论的根本原因，无非是人的认识与客观实际以及认识客观世界的方法与客观规律的矛盾。这种直接和间接的矛盾在某一点上的集中表现就是悖论。例如，伽利略悖论和希帕索斯悖论就表现为直接的主观认识与客观实际的矛盾。又如康托用于造集的概括原则是认识客观世界的方法或手段，但是生成集合有它自身的客观法则，那么产生罗素悖论的原因乃在于概括原则造集的任意性与生成集合的客观规则的非任意性之间的矛

盾，这就表现为认识世界的方法与客观规律之间的矛盾，这是一种间接的主客观矛盾，一定历史阶段中人的主观认识与客观实际之间的矛盾．而悖论就是这种矛盾在某些点上的集中表现。

如此，由于人的认识在各个历史阶段中的局限性和相对性．在人类认识的各个历史阶段所形成的各个理论体系中，本来就具有产生悖论的可能性，但在人类认识世界的深化过程中同样具备排除悖论的可能性和现实性，人类认识世界的深化没有终结，悖论的产生和排除也没有终结。因之，在绝对意义下去寻求什么产生悖论的终极原因和创造什么解决悖论的终极方法都是不符合认识论原则的。

悖论问题的研究，对于数学基础理论、逻辑学、语言学和哲学的研究都有重要意义，试看语义学、类型论、多值逻辑、公理化集合论的几个重要系统，直到近代数学的三大流派的形成和发展，无不直接来自悖论问题的研究，还有公理化方法论，数理逻辑、证明论和模型论的形成和发展的一个切近原因，除了非欧几何的产生之外，仍然是悖论问题的研究。特别是我们在§6中讨论的哥德尔不完备性定理这样的重大研究成果，其直观背景和证明思想也直接来自悖论的分析。所以，贝特说："现代逻辑中的许多最为深刻的成果，都从悖论的分析中产生"。Fraenkel说："作为近代基础理论研究的最有兴趣的发展之一，就是证明了这样一点，语义学悖论所代表的问题不仅是那种与数学本身只有间接关系的方法论问题，而且也是对于数学有着巨大直接作用的研究的出发点"。塔斯基说："必须强调的是，悖论在建立现代演绎科学的基础占有一个特别重要的地位，正象集合论的悖论，特别是罗素悖论成为逻辑和数学相容性形式化的起点一样，撒谎者悖论及其他语义学悖论导致了理论语义学的发展"。总而言之，在二十世纪的今天来讨论和研究悖论问题，首先应该把十八世纪以前那种认为悖论只是茶余酒后的闲谈的陈旧看法扫除，否则只能说明他对数学基础、数理哲学的近代发展视而不见。

第9讲　论数学基础诸流派及其无穷观

§1　数学系统的相对相容性证明与诸流派形成的历史近因

所谓数学系统的相对相容性证明就是把一个公理系统的相容性证明化归为另一个看上去比较可靠的公理系统的相容性证明，也即依靠某一个数学系统的无矛盾性来保证另一个数学系统的协调性。这一概念的产生和历史发展的背景是这样的：自从Н.И.Лобачевский几何诞生后，由于罗氏平行公理如此地为常识所不容，这才真正激起了人们对于数学系统的无矛盾性证明的兴趣和重视。后来，庞卡莱在欧氏半平面上构作了罗氏几何的模型，这就把罗氏系统的相容性证明通过一个模型化归为欧氏系统的相容性证明，但却由此而导至了人们对欧氏系统相容性的重重疑虑幸亏那时已经有了解析几何，这就等于在实数系统中构造了一个欧氏几何的模型。这就把欧氏几何的无矛盾性归约到了实数论的相容性。那么实数论的相容性如何？Dedekind把实数定义为有理数的分划，也即有理数的无穷集合。因而把这个无矛盾性归约到了自然数系统的无矛盾性。又由于Frege和Dedekind的自然数概念是借助集合的概念加以定义的。因此，归来归去还是把矛盾集中到集合论那里去了。那么集合论的相容性如何？事实上，集合论的相容性正处于严重的"危机"中，以致这种相对相容性的证明至今还是一场空。

在非欧几何获致普遍接受及其相对相容性得到证明之前的很长历史阶段中，人们几乎确信数学真理就是绝对真理，例如Kant就把欧氏几何看成是关于空间的绝对真理，即所谓先验的综合

判断。然而，上述庞卡莱模型的出现，特别是人们相继发现欧氏几何的每条公理在罗氏空间的极限球上得以全部成立，从而欧氏几何的相容性反过来也可借助非欧几何的相容性加以保证。这样，平行公理完全相背的两个几何系统竟然是互为相容的，这就直接否定了上述数学真理就是绝对真理的观点。待至各种几何系统在空间曲率概念和群论观点下获至统一的时候，人们则更是感到这样多的至少是部分地矛盾的几何居然都能用来描述物理空间，我们真不知道，对于物理空间来说，究竟哪一种是真实的了。在部分数学家那里，竟导致了对于数学真理性的否定而走向了相对主义。

此外，集合论悖论的出现，震动了整个西方数学界和逻辑学界，希尔伯特指出："必须承认，由于悖论的出现而造成的形势是难以忍受的。只要设想一下，每个人曾学过、教过并在数学中加以应用的定义和演绎的方法，从来都被认为是真理和必然的典范，而现在却导致了荒谬，如果连数学思维都是不可靠的，那么到哪里还能找到真理和必然性呢？"

数学是演绎推理性质的学科，所以从形式上看，数学命题的真理性是建立在公理的真理性和逻辑规则的有效性之上的，本来即使在上述"数学真理是绝对真理"这一观念受到冲击之后，数学家还可以用逻辑推理的严格性来作为精神支柱，并以此解释数学在应用中的有效性，但由于悖论的出现，这一精神支柱也动摇了。这就不能不在数学家中形成了"危机"感。

正因为数学面临着这样的"危机"，才促使数学家们去探索数学推理在什么情况下有效，什么情况下无效，数学命题在怎样的情况下具有真理性，在怎样的情况下失灵。于是，在本世纪初，数学基础论这一分科就诞生了。摆在从事数学基础问题研究的数学家面前的首要任务，就是如何为数学的有效性重新建立可靠的依据。由于在这一工作中所持的基本观点不同，才在数学基础论的研究中形成了各个流派。而这也正是形成数学基础诸流派

的共同的历史背景。

§2 逻辑主义派的观点和方法

逻辑主义派的主要代表人物是罗素。逻辑主义派的主要宗旨是把数学化归为逻辑。亦就是说：第一，数学的概念可以从逻辑的概念出发，经由明显的定义而得出；第二，数学的定理可以从逻辑的命题出发，经由纯逻辑的演绎推理而得出。因此，全部数学都可以从基本的逻辑概念和逻辑规则推导出来。这样一来，数学也就成为逻辑的分支了。

逻辑主义的形成，不仅有它的历史背景，还有它自身的历史发展过程。其实，逻辑主义之思想原则的萌芽，亦即把逻辑看成是先于一切科学的观点，可以追溯到Leibniz，但他本人并没有从事这方面的工作，他的这一想法直到十九世纪才在Dedekind、Frege、Peano等人的工作中得到初步发挥。应该说，逻辑主义的观点在Frege的工作中就基本形成。因为Frege明确地提出了数学可以化归为逻辑的想法，而且花费了将近二十年的时间把算术化归为逻辑，这就是他的巨著《算术基础》和《算术基本原理》。另外，Dedekind也应该是逻辑主义的创始人之一，因为他在1887年发表的《数的性质与意义》一文中就明确提出了与罗素几乎完全相同的主张。当然，Frege的工作没有超出算术的范围，Dedekind对自己的所述主张也没有更多的展开，只有在罗素和Whitehead那里，才把这派的主张详详细细地加以发展，并且真正从相当少的公理和概念出发，导出了大部分数学。

为了弄清罗素的数学观，不妨先看一看罗素是沿着怎样的道路去从事数学基础理论的研究工作的。对此，罗素在他的《数理哲学引论》一书中指出："数学是这样的一种研究，它可以按两个方向去进行。一个比较熟悉的方向是建设性的，即不断增大理论的复杂性，……，另一个是不很熟悉的方向，这就是通过分析来

达到越来越大的抽象性和逻辑简单性，这里所考虑的已不再是从怎样的假设出发可以定义或演绎出什么结果的问题，而是研究我们能否找到更为一般的思想原则，从这些思想和原则出发，能使现在作为出发点的东西得以被定义或演绎出来。”罗素在数学基础理论上的研究工作是沿着后面这个不很熟悉的方向进行的。那么，罗素后来所找到的这些更为一般的思想和原则是什么呢？这就是罗素所指出的："应当从一些普遍承认是属于逻辑的前提出发，再经过演绎而达到那些明显地属于数学的结果。”而这也正是实现逻辑主义根本宗旨——把数学化归为逻辑——的根本做法。

从逻辑主义的形成和历史发展来看，逻辑主义的工作可以划分为数学理论的算术化和算术理论的逻辑化这样两个阶段。

十九世纪的最后二十五年，堪称为数学理论算术化的时期，许多出色的数学家都投入了这个工作。不仅数学分析建基于实数论，几何也归约到了实数论。1872年Weierstrass、Cantor和Dedekind等人几乎同时完成了实数定义，但无论是康托的收敛有理数序列还是Dedekind的有理数分划，都表明了实数论被化归为有理数论，进而整数直至自然数系统，当然其中已借用了集合论概念。从而世所著称的、由三个原始概念和五条公理所构成的皮亚诺(Peano)算术公理系统便成为这个数学理论算术化工作的终结。人们从皮亚诺系统出发，借助集合论的概念，便可建造算术、分析、几何直至整个数学大厦。因而这种数学理论算术化使得许多数学家感到满意，甚至认为无需再前进了。但对逻辑主义者来说，这只是实现逻辑主义宗旨的第一步，要把数学化归为逻辑，更重要的是所谓算术理论逻辑化。

Frege、Whitehead和罗素曾为借助于纯逻辑概念去定义皮亚诺系统的三个原始概念、借助于纯逻辑法则而演绎地引出皮亚诺系统的五条公理等等付出了巨大的劳动。后来，罗素曾经认为，这种算术理论逻辑化的工作已经完成，从而认为逻辑主义的宗旨已经实现。这就是罗素自己所说的："我们发现在这里(即指《数学原

理》一书——引注）没有一个地方能作出明确的分界线，使得逻辑在其左方，数学在其右方，如果还有人不承认逻辑和数学的同一性，我们就请他来证明，《数学原理》中的哪些定义和演绎可看成是逻辑的终点和数学的起点。”“从逻辑中展开纯数学的工作，已经由Whitehead和我在《数学原理》一书中详细地做了出来。”其实，事情并没有这么简单，经过仔细分析，普遍认为，罗素和他的合作者并没有能在《数学原理》一书中实现逻辑主义的宗旨，以上的说法，只能是罗素的一个自我安慰而已。

事实上，人们在《数学原理》一书中可以清楚地看到，在他们从“逻辑”出发推导数学命题的过程中使用了集合论的公理，特别是使用了无穷公理和选择公理。因为没有无穷公理，连自然数系统都无法构造，更谈不上全部数学了。人们现已引进了所谓初等公理系统与高等公理系统的概念，凡是无需借助无穷公理的系统称为初等的。凡是必需借助无穷公理的系统称为高等的。而算术与几何理论均属高等公理系统，亦即不可能不借助于无穷公理。其次，没有选择公理，则数学中的定理就要砍掉一大批。罗素和Whitehead也在事实上借助了这两条公理。那么，这两条公理究竟算不算逻辑公理？这首先又要牵涉到什么是逻辑的问题，按照人们对于逻辑的一般理解，普遍认为纯粹逻辑只涉及形式而不涉及具体事实，亦即作为逻辑法则，只允许讨论可能性对象，而不允许在逻辑法则中作出某物是否存在的断言。但无穷公理恰恰就是一条存在性公理，因为存在一个集合是陈述公理的必然判断，选择公理亦复如此，因之普遍认为他们不是逻辑公理。既然出手就借助了非逻辑公理，怎能说全部数学命题均由逻辑规则和逻辑概念演绎推出的呢？所以Frege在晚年已倾向于放弃逻辑主义的立场。至于罗素自己，其实心里也明白，因为他们在使用无穷公理时显得很勉强，正如Fraenkel和Bar-Hillel所说：“他们是很勉强地走这一步的。”“Frege和罗素的理论的严重的缺陷在于无穷公理的令人怀疑的状况。”王浩也指出：“正是由于要借助

严格的逻辑概念来给出'无限'的一个充分根据这一困难，才使罗素关于数学与逻辑相等价的理论成为可疑。”因此，明摆的事实迫使罗素和Whitehead对此设法补救，其补救办法是把所有要用到这两条公理之一才能证明的命题一律改写为条件式命题，亦即不把这两条公理列入他们的系统，但当某数学命题P必须借助无穷公理才能证明时，就把P陈述为“如果无穷公理真则P”等等，因而这两条公理就变成诸如此类的每一定理的前提，而不再是特设的公理了。又Carnap曾建议采用“坐标语言”而废除“名称语言”的办法补救之，在他看来，只要这样做了，则无穷公理就可以被认为是关于“位置”而不是关于事实的断言，而位置则可能是空的。但是，不论是罗素的条件式命题还是Carnap的坐标语言，普遍认为这都是勉强的做法，也就难以取得大多数数学家的支持了。只有罗素等少数人才认为已经借此实现了逻辑主义的宗旨。

另方面，问题的复杂性还远不止于此。大家知道，要从逻辑导出全部数学，势必要导至无限集理论的展开，而集合论本身矛盾重重的盖子正是以著名的罗素悖论揭开的，而逻辑系统是不允许有矛盾的，因之集合论的悖论必需排除。罗素借以排除集合论悖论的办法是引进“恶性循环原则”而发展他的分支类型论。但是这个分支类型论不仅显得这样复杂而支离破碎，更严重的是由此就不能实现逻辑主义之宗旨。因为恶性循环原则是直接排斥非直谓定义法的，而非直谓定义法的禁止使用就势必要抛弃许多有用的数学概念和命题。因此，摆在罗素面前有两条路，或者放弃恶性循环原则，从而也就拆除了防止悖论泛滥的罗素式的堤防，或者坚持恶性循环原则，从而也就无法实现逻辑主义的主张。无疑地，这两条道路罗素是一条也不愿意走的。他终于在无路可走的情况下找出了一条虚假的小路，那就是引进可化归公理：任何（广义）公式都可以和一个直谓（广义）公式相等价。但是，如果我们以公理的形式肯定了任何一个非直谓（广义）公式总可有

一个等价的直谓（广义）公式来取代的话，则任一类中较高级的性质就可化归为同一类中较低级的性质，以此类推，便把类中之分级全部取消了。从而剩下的就是没有恶性循环原则的简单类型论了。所以，人们稍加分析，便发现可化归公理的精神与恶性循环原则相冲突。引入可化归公理后，实质上便是取消恶性循环原则。因而人们议论纷纷，可化归公理一时成为人们批评《数学原理》之众矢之的。普遍认为，可化归公理过于人为而不自明，其作用无非是把分支类型论约化为Ramsey的简单类型论，所以大多数人宁可直接采用简单类型论而不愿意在可化归公理的虚假小道上绕弯路。罗素也终于放弃可化归公理，但又抓住恶性循环原则不放，从而在实际上放弃了逻辑主义的主张。但罗素是直到晚年才承认："我所一直寻找的数学中的光辉的确定性在令人困惑的迷宫中丧失了。""寻求完美、最终和确定性的希望破灭了。"

综上所述，我们已经看到，逻辑主义宗旨之实现不仅困难重重，罗素的方案也实在问题百出，只好以失败告终。其实，罗素方案失败的根本原因在于只看到并且过于夸张了数学与逻辑在演绎结构上的同一性，而完全抹杀了数学与逻辑科学的质的差异性。但是，我们却不能由于逻辑主义的失败而认为逻辑主义者的工作一无是处或毫无价值。相反地，应该看到，逻辑主义者的工作实际上已对数学与逻辑的发展作出了重要贡献：

（1）众所周知，罗素的分支类型论，特别是经过Ramsey改进之后发展起来的简单类型论，对于悖论的研究和排除是有重要意义的。而且现有的一些解决悖论的方案，看来无不渊源于罗素早年提出的见解。

（2）逻辑主义者已相当成功地把古典数学纳入了一个统一的公理系统，虽然这个系统不是纯逻辑的，但这样一个工作却成为公理化方法在近代发展中的一个重要起点（参阅《数学原理》）。

（3）由于逻辑主义者的工作，基本上完成了从传统逻辑到数理逻辑的过渡和演变。特别是Frege、Peirce、Schröder等一

起最早地引进了量词，并对量词的性质作了深刻的研究。Frege还进一步给出了命题演算和谓词演算的系统。

最后，让我们简述一下逻辑主义派的无穷观，由于逻辑主义派的基本立场是确认全部数学的有效性，并认为能把全部数学化归为逻辑，因此，既要确认全部数学的有效性，势必要确认实无限观点下的无限集理论。因此，就无限观而言，逻辑主义派是实无限论者，亦即确认实无限性研究对象在数学领域中的合理性。普遍认为罗素及其追随者明显地承认无限性对象的存在性。但由于罗素为排除集合论的悖论而发展他的分支类型论，从而在罗素系统中的实无限性对象就在不同的类和级中表现为一定的层次结构。这是符合反映论者的见解的。

§3　直觉主义派的观点和方法

直觉主义派的主要代表人物是Brouwer。直觉主义派的根本出发点是关于数学概念和方法的"可信性"考虑.因此,认识论上的可信性就唯一地决定了直觉主义的前提。直觉主义者Heyting说："当你们通过公理和演绎进行思维时，我们则借助'可信性'进行思维，这就是全部的区别。"

直觉主义者认为,集合论悖论的出现不是一个偶然事件,它是整个数学所感染的疾病的一个症兆，因此，悖论问题不可能通过对已有数学作某些技术性的修改或限制而得以解决，必须依据可信性的要求对已有的数学作全面审查，而且应该毫不犹豫地放弃那些不符合可信性要求的数学概念和方法.那么,什么样的数学概念和方法才算是符合直觉主义的可信性标准呢？这个标准就是直觉主义的著名口号："存在必须是被构造。"亦即数学中的概念和方法都必须是构造性的。所谓主张概念和方法上的构造性，就是只承认按固定方式经有限个步骤能够定义的概念和能够实现的方法才是有效的，故构造性亦称能行性，构造性的方法亦称能行

性的方法。例如，求两个正整数a，b的最大公约数，可用Euclid除法在有限步骤内实现。象这类方法就称为能行的或构造性的方法。

直觉主义者在数学上的出发点不是集合论，而是自然数论。这就是Heyting所说的："数学开始于自然数及自然数相等概念形成之后。"围绕着自然数论的可信性问题，直觉主义者指出，自然数来源于Brouwer的"原始直觉"或称"对象对偶直觉。"所谓对象对偶直觉，即所谓人皆有之的一种能力——某一时刻集中注意某一对象，紧接着又集中注意于另一对象，这就形成了一个原始对偶，就用（1，2）来表示它，有了这个原始的对象对偶，便可根据构造性的要求重复一次而产生（2，3），再重复一次便是（3，4），依此递推下去，则任何一个自然数都能从这个对象对偶直觉开始，用构造性的方法产生出来。直觉主义者认为，只有建立在这种原始直觉和可构造之上的数学才是可信的，而这种原始直觉对于思想来说是如此直接，其结果又是如此清楚，以致不再需要任何别的什么基础。

应当指出，不要把直觉主义者的"直觉"与马克思主义认识论中所论述的"直观感觉"混为一谈。相反地，在哲学观点上，直觉主义派是彻底的主观唯心论者。Heyting明确指出："如果把直觉主义的断言看成是关于事实的断言，那将是一种武断，因为它并不具有这种意义。""直觉主义的推论不是关于事实的推论，而是一种理智的构造。""我的数学思想属于我个人的理智生活，并限于我个人的思想……数学思想的特性是在于它并不传达外部世界的真理，而只与心智的构造有关。"所以直觉主义者的所谓"原始直觉"和由此开始的"构造"，用他们自己的话来说乃是一种内省直觉能力的发挥。

直觉主义派也有它本身的历史发展过程，如果单纯着眼于无穷观，则直觉主义派的思想一直可以追溯到古代，因为从"存在必须被构造"这一前提出发，势必导致彻底的潜无限观念。历史上第一次明确地只承认潜无限而反对实无限的是Aristotle。略

早于他而明确承认实无限的是Plato。Aristotle的基本倾向是在依靠逻辑的同时，更大力地依靠感官的直接印象，而对于感觉能力所不能及的抽象和外推是不喜欢的。这就势必要对实无限采取排斥的态度。他也明确认为，无限只能是一种潜在的存在，而不能是一种实在的存在。Aristotle这种无穷观也是往后数学领域中一切潜无限论者的基本观点。

如果立足于直觉主义派的总的哲学观点，则Brouwer自已指出："我们可以在Kant那里找到直觉主义的一种古老的形式。"但Brouwer却自称为新直觉主义者，因为他放弃了Kant关于空间的先验性，而更坚持关于时间的绝对先验性。事实上，Brouwer所说的原始直觉，就是Kant关于时间的直觉，而Brouwer关于自然数渊源于原始直觉的提法也就是Kant关于自然数是从时间直觉中推演出来的主张。但在Kant那里，除掉时间直觉以外，还有关于空间的直觉，几何理论则产生于空间直觉。在Brouwer那里，除了原始直觉以外，不再需要任何别的纯粹直观。因为一旦有了这个原始直觉，往后的一切就可以从此开始去构造了。

直觉主义派在数学上的直接先驱者乃是与康托同时代的Kronecker，因为他明确提出并强调了能行性，主张没有能行性就不得承认它的存在性。在无穷观的问题上他又是实无限概念的激烈抨击者，他与康托进行了长期的针锋相对的争论。他曾计划要把数学算术化并在数学领域中清除一切非构造性的成分及其根源。庞卡莱也是直觉主义在数学上的先驱者，因为他主张自然数是最基本的直观，无需再作进一步的分析就可以认为是可信的。其次，庞卡莱也曾多次谴责完成了的无穷集合观念，并主张潜无限观念，他也不赞成形式化的研究方法。但是庞卡莱却始终没有把这种倾向明确地加以总结并上升为基本观点。另外，法国的半直觉主义者也为直觉主义的形成作了直接的准备，因为他们在抨击选择公理的同时强调了能行性。当然，只有在Brouwer那里，才完整而彻底地从哲学和数学两方面贯彻和发展了"存在必

须被构造”的观点。所以大家就公推他为直觉主义的奠基者和代表人物了。这一学派中的主要人物还有Heyting和Weyl等人。

直觉主义派的基本观点直接决定了这一学派在数学工作中的基本立场是：第一，在无穷观的问题上彻底采纳潜无限而排斥实无限；第二，否认传统逻辑的普遍有效性而重建直觉主义逻辑规则；第三，批判古典数学，拆除一切非构造性数学的框架，重建直觉主义的构造性数学。下面我们来作进一步的讨论和评述。

（一）直觉主义派的无穷观：根据直觉主义的基本观点，势必导致对实无限概念的排斥。因为从生成的观点来看任何一个无穷集合或实无限对象都是不可构造的。若以最简单的自然数集为例讨论的话，按照能行性的要求必然否定自然数全体这个概念，因为任何有穷多个步骤都不能把所有的自然数构造出来，更谈不上汇成整体了。而且即使先假设有那么一个全体自然数论域摆在那里的话，直觉主义者也不承认能把全体自然数逐一复查完毕，亦即不承认有走遍自然数论域的概念。在他们看来，自然数1，2，3，…，只能永远处于不断地被构造的延伸状态中。例如，直觉主义者Weyl明确指出：“Brouwer使这一点明确了，就是没有任何证据能够证明所有自然数的整体的存在性，……，自然数列，它能够通过不断地达到下一个数而超越任何一个已经达到的界线。从而也就开辟了通向无限的可能性。但它永远停留于创造（生成）的状态之中，而绝不是一个存在于自身之中的事物的封闭领域。”由此可见，在无穷观的问题上，直觉主义派是十分彻底地采纳了潜无限论者的观点。如所知，实无限论者对自然数全体的认识和理解是完全不同的，不仅完全肯定自然数全体这一概念，并据反映论观点作出实际的解释。实无限论的反映派认为，人类对自然数无穷序列的认识是经过了几个不同等级的抽象才完成的：第一是由具体事物到自然数概念，这是一级抽象。第二是由具体的自然数到一般的自然数n，这是二级抽象。第三是从任意有限多个自然数到自然数全体，这是三级抽象。而人们之所

以能完成这第三级的抽象过程，主要是因为思维能够反映事物在质变过程中的“飞跃”。在这里便是具体反映了从延伸到穷竭、有限到无限或限量到质的转化。借助于如下的实际考察，对于自然数全体的概念是很直观的。如图设某物P沿箭头所示方向以每秒

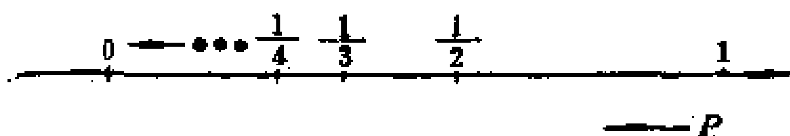


图9-1

一个单位的速度从1处向0处运动。这从现实的位移运动观点来看，P是完全可以实现从1到0的位移运动的，它首先以$\frac{1}{2}$秒时间从1移到$\frac{1}{2}$处，在$\frac{2}{3}$秒时间后就经过了$\frac{1}{3}$处，而在$\frac{3}{4}$秒时间后经过了$\frac{1}{4}$处，以此类推，总共用了1秒钟时间由1移到了0。而P在这一秒钟时间里就经过了$\frac{1}{1}$，$\frac{1}{2}$，$\frac{1}{3}$，…，$\frac{1}{n}$，…所有各个点，其分母也就走遍了全体自然数而使之汇成一个整体。所以从实无限的反映派观点来看，自然数全体这一概念是能够接受的，其中关键的一步就是承认思维对延伸到穷竭这一飞跃的正确反映。但从构造性观点出发，就必然否认思维对于飞跃的能动反映。在这一点上，大致所有的潜无限论者都是如此。

（二）直觉主义派的逻辑规则：出于存在必须是被构造的考虑，直觉主义派否认传统逻辑的普遍有效性。Brouwer认为古典逻辑是从有穷集合抽象出来的，不能无限制地使用到无限性对象上去，悖论就出在无穷上。Brouwer指出：“人们把逻辑误认为是某种超越和先于全部数学的东西，并且不加检验地把它应用到关于无穷集合的数学。”Brouwer在这里说了两件事：第一，传统的逻辑规则不应该无条件地应用于无限性论域上；第二，认为逻辑先于全部数学则是一种误解。所以，直觉主

义正好与逻辑主义相反，非但不把逻辑看成是先于数学而为数学之基础，反过来视逻辑为数学的一个部分，直觉主义者进一步指出："逻辑并不是我们站立的基地……事实上，它不过是一种具有特殊的一般性的数学定理，亦即逻辑只是数学的一个部分，而决不能作为数学的基础。"因此，直觉主义派也发展了自已的逻辑，这是和传统逻辑很不相同的一种逻辑。一般地说，这种不同之处主要表现在否定性质的推理上。现把一些不同之处列举如下：

（A）排中律：设S表示一个无穷域，如前所提及，实无限论者认为S是一个既经完成了的无穷总体。因之，建立在实无限观点上的古典数学便认为S是一个已经构造完成的封闭域，从而能把S的全部个体逐一检查完毕，亦就是我们在推理中常说的"走遍S的一切元等等。"但是直觉主义派既然认为任何无穷集合都只处在无止境的构造中而不是已经构造完成了的封闭域，从而决不承认能把所有某类具有性质P的个体检查完毕。因而直觉主义派就不允许把排中律$P\vee\neg P$使用到无穷集合上，亦即只承认排中律在有穷集合上的有效性。例如，直觉主义者认为能对$\{1, 2, 3, \cdots, n\}$这n个自然数下结论："每个自然数或者是偶数，或者不是偶数。"但不能对所有的自然数下这个结论。因为对有限多个自然数而言，我们能给出一个能行的过程在有限步骤内把这有限多个自然数一一检查完毕，但对全体自然数来说，按照构造性观点只能做到哪里算到哪里，下一步情况如何，等做出来之后再说，在下一步还没有检查之前，我们是不能对它作什么结论的。由于这种一步一步的检查永远做不完，所以也不会存在一个能行的过程能在有限步骤内把所有的自然数遍查一次，以致能对所有自然数判定其每一自然数要么是偶数、要么不是偶数。所以，根据构造性的观点，把排中律应用于所有自然数之上是无效的。

直觉主义者又明确指出，所谓一个命题是真的，就是说存在一个能行的过程在有限步骤内已经证明了该命题是真的，一个命

题是假的，也是指已经能行地证明了该命题为假。然而事实上，存在着大量的数学命题既没有能行地证其为真，也未证其为假。这些未证命题是不是时间未到，今后时间一到就总能或证其为真或证其为假，这一点直觉主义认为根本不知道。但是未证命题一直是越来越多，而不是越来越少。一个多年未决的难题的解决往往带来一批未证命题。因此，直觉主义更认为，排中律“命题A或命题$\neg A$必有一真”是不能承认的。因为根据构造性的观点只能是证一个算一个，证到哪里算到哪里，如果无条件承认命题A或命题$\neg A$必有一真的话，这就等于承认了所有命题总是能构造性地证其为真或不真的。但在直觉主义者看来，这是没有可信性根据的。这就是Brouwer所说的："承认排中律实际上就等于承认对每个数学命题都能或证其为真或证其为假。”Heyting也指出："由Brouwer的基本观点，即研究心智的数学构造而不涉及被构造对象的性质，就立即导致了对排中律的拒绝。”但在古典数学中，由于承认排中律，所以一个命题不论能不能或是不是已证其真或假，该命题总归是或者为真或者为假，直觉主义则不然。

(B) 反证律与非古典逻辑演算：直觉主义者认为，欲证明命题A为假，当可设A真，然后给出一个能行的过程在有限步内引出矛盾，从而结论A假。亦即在能行性要求下承认如下的形式推理规则：

($\neg_+$) 如果Γ，$A\vdash B$，$\neg B$，则$\Gamma\vdash\neg A$。

但是对于一个命题的肯定式，使用反证法来证其为真是不为直觉主义者所接受，亦即虽然在能行的要求下也不承认如下的形式推理规则：

($\neg_-$) 如果Γ，$\neg A\vdash B$，$\neg B$，则$\Gamma\vdash A$。

通常称($\neg_-$)为反证律，它反映演绎推理中的反证法。通常称($\neg_+$)为归谬律，它反映演绎推理中的归谬法。直觉主义者认为象($\neg_-$)这样的推理规则太强，不能无条件地承认。其实根本问题还是由于构造性观点与无条件承认排中律是不相容的。

(C) 量词的解释与等价式：由于直觉主义者强调能行性，对于量词的解释也不同于古典意义下的理解。例如，$\exists XA(X)$虽然也被解释为"有一个X使$A(X)$成立"，但基于构造性观点，就必须给出一个能行的过程在有限步骤内把那个使$A(X)$成立的X找出来。否则不能算是在直觉主义观点下的"有X使$A(X)$成立"。如此，直觉主义对于"如果$\neg\exists XA(X)\vdash B$，$\neg B$则$\vdash\exists XA(X)$"是根本不承认的。因为仅由"若没有$X$使$A(X)$成立而导出了矛盾"这一点，并不能说明已经给出能行地找出那个使$A(X)$成立的X的方法。又如一些古典谓词逻辑中有定理：

$$\neg\forall XA(X)\vdash\exists X\neg A(X).$$

直觉主义者认为不是普遍有效的。因为直觉主义者不承认能对无限论域S之所有个体能能行地一一检查完毕。所以"不是所有的X使$A(X)$成立"这句话对于能行意义下的"有一个使$\neg A(X)$成立"这句话来讲乃是一句空话，因为单凭前一句话是不能给出能行地找出$\neg A(X)$成立的X的方法的。如果你是在已经给出了能行地找出使$\neg A(X)$成立的X的方法之后才说"$\neg\forall XA(X)$"这句话的话，则直觉主义者就要请你说"$\exists XA(X)$"而不要画蛇添足地去多说"$\neg\forall XA(X)$"这样的废话了。所以直觉主义者认为如上的定理要么是无效的，要么是多余的，因而是不接受的。所以在古典系统里承认等价式：

$$\neg\forall XA(X)\dashv\vdash\exists X\neg A(X)$$

但直觉主义只承认$\exists X\neg A(X)\vdash\neg\forall XA(X)$而不承认$\neg\forall XA(X)\vdash\exists X\neg A(X)$.

（三）直觉主义派的构造性数学：基于构造性观点，势必要排斥古典数学中的非构造性数学。让我们先举一例，借以对照说明非构造性观点与构造性观点处理问题时的不同思想方法：

例如，让我们考虑圆周率π的十进位小数表示式

$$\pi=3.14159265358979\cdots,$$

令$f(n)$表示第n位小数前出现数字5的个数，如$f(3)=0$，$f(4)=1$，$f(8)=2$，$f(10)=3\cdots$等等。

试问不等式$\frac{f(n)}{n}\leqslant\frac{1}{2}$是否对每个自然数$n$都成立?对此问题，古典数学认为答案或者是肯定的，或者是否定的，两者必居其一。直觉主义数学则认为根本不能回答。

古典数学对这个问题的处理方法是：

1° 因为π有一个确定的解析表达式，因此π的十进无尽小数表示式是存在而有意义的。

2° $\left\{\frac{f(n)}{n}\right\}$作成一个有界的无穷数集，其中$0\leqslant\frac{f(n)}{n}<1$，而$n=1, 2, 3, \cdots$。

3° 根据Dedekind割切原理，有界数集$\left\{\frac{f(n)}{n}\right\}$必存在确定的上确界$\sup\left\{\frac{f(n)}{n}\right\}=\beta$，$\beta\leqslant 1$。

4° 将β与$\frac{1}{2}$作一比较，根据三分律可知，$\beta>\frac{1}{2}$与$\beta\leqslant\frac{1}{2}$二关系式中有且仅有一个关系成立。

5° 因此这个问题的答案或是肯定的，或是否定的，两者必居其一。即当$\beta\leqslant\frac{1}{2}$，则肯定，当$\beta>\frac{1}{2}$，则否定。

按照直觉主义的思想方法，则对于以上推理过程是根本不能接受的。首先对于π被表示为十进无尽小数展开式是无法考虑的，只能算到哪一位就算是展开到哪一位。其次是不承认$\left\{\frac{f(n)}{n}\right\}$是一个完成了的有界数集，因为自然数$n$都只能在永无止境的被构造之中，数$\left\{\frac{f(n)}{n}\right\}$也只能在无止境的被构造之中，对于$\sup\left\{\frac{f(n)}{n}\right\}=\beta$的存在性更是不承认的。按照构造性观点，对于$\frac{f(n)}{n}$的验算也只能算到哪一步把话说到哪一步，所以，对于所有的自然数来

问$\frac{f(n)}{n}\leqslant\frac{1}{2}$是否成立？直觉主义派拒绝回答。

又如考虑连续函数性质的中间值定理："设$f(x)$在$[a, b]$上连续，并且$f(a)\cdot f(b)<0$，则存在一个c $(a<c<b)$，使得$f(c)=0$。"我们知道在数学分析中是用反证法来证明这个c的存在性的，但并没有给出一个能行的过程在有限步骤内确立这个c的存在。自然，直觉主义者不接受这种非构造性证明。

今以直觉主义分析学为例，试看直觉主义者如何建立直觉主义数学。如所知，古典分析建基于实数连续统，因此，建造直觉主义分析学的根本问题就在于如何在可构造意义下给出实数和实数连续统的概念。因此，直觉主义者首先引进所谓"属种"(spècies)的概念以取代康托意义下的集合概念。例如，只要能给出一组确切的、能为有限句逻辑上无矛盾的语句所表述的规则L，根据L就能把每一自然数一个接一个地、无止境地构造出来，这就算是给出了一个全体自然数的属种。进一步，Brouwer又引进了"选择序列"的概念："在任何时刻，一个选择序列α系由一个有穷的节连同对它的延伸的若干限制组成。"如此，直觉主义者便以"有理数选择序列"取代古典分析中的有理数柯希序列概念，并称之曰："实数生成子。"相应于古典分析中把实数定义为有理数柯希序列等价类，可构造意义下的单个实数被定义为实数生成子的一个等价属种。如上所见，建立可构造性实数概念没有实质性困难，其原因就在于Cauchy－Weierstrass的整个极限论建基于潜无限观念。因而在实质上，直觉主义者在此不过是在能行性的要求下重新陈述柯希序列而已。就是说，在相应的陈述下，对于任给$\varepsilon>0$，必须总能够能行地找出一个$\delta>0$，而决不是非构造性地存在一个$\delta>0$等等。

建造直觉主义分析学的真正困难还在于"构造性连续统"概念的建立。其原因在于可构造性至多只有潜无限，因而无限至多只有一个层次，但实数却有不可数多个，对于可数无穷来说乃是两个

不同层次的无穷。因而问题不单是以潜无限取代实无限，而要把高层次的无穷统一到低层次的无穷上来。Heyting对此指出：“如所周知，递归实数没有穷尽连续统，递归实数是可数而连续统是不可数的。Brouwer曾尝试找一个尽可能与普通的连续统接近的构造性的概念。他为这个问题奋斗终身。在他的1907年学位论文中，他引入连续统作为初始概念，认为人有连续统的一个直观（时间的直观），靠它他能构造一稠密的、可数无穷的标尺，连续统上的一个点是用这标尺的点的一收敛序列定义的。”但是，Brouwer一旦借助于时间的直观而提出连续统也是原始直觉的说法，就显然违背了直觉主义者的基本立场，以致陷入自相矛盾的境地。直到1919年，Brouwer终于利用“展形”（spread）概念巧妙地建造了符合构造性要求的连续统概念。其中关键的一步就在于不再是先实数后连续统，而是把每一个实数同时统一在一个潜无限的构造性状态之中。为使直观图象简单明白起见，我们把展形连续统意译为如下一个通俗易懂的展形图象，并在解释中采用二进位记数法。

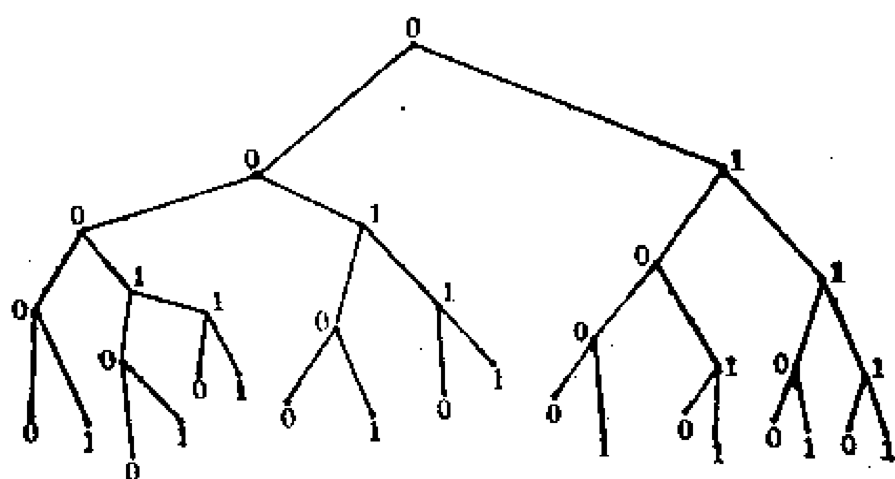


图9-2

不妨把上图所示的这个展形连续统理解为一棵永远在生长着的树，其生长的规则是先由树根长出两枝，每枝长到一有限长时

又各生出两枝，以此类推，永无止境地长下去，每个开始长出分枝之处称为“节”，又把每个“节”生出的两枝的“节”合在一起称为一“节对”，于是除去树根的那个节之外，每个节都在一个节对中，现把每个节对中的一个节记为0，另一个记为1，不允许同时记为0或同时记为1，0与1称为节的标记。这样，从任何一个节到它紧接的下一个节就有且仅有0或1两种选择的可能。我们把树根以下的一个节开始一个接着一个节生长下去的变程称为此展形树的一个分支。如果把任一分支中从第一个节的标记起依次记录下来便是一个可构造意义下的实数二进表示式。反过来，任给一个可构造意义下的实数二进表示式，那么按照上述方法必能在此展形树上找出一个分支，它的依次相接的节的标记构成这个可构造实数二进表示式。在这里要注意的是，可构造实数二进表示式的位数虽然可以无止境地增多，但却总是有限的，而且只能给到哪一位算到哪一位，因而完全不同于古典分析中实无限意义下的实数二进表示式的含义。

如此，直觉主义意义下的单个实数就既不是在连续统生成之前，也不是在连续统生成之后一个一个地被构造出来，而是在建造出实数连续统的同时构造了每一个实数。反过来，构造性意义下的实数连续统也是在构造出每一个构造性实数的同时被构造出来，从而直觉主义意义下的连续统本身与每个实数同时处在能行的、潜在无限的被构造状态中。这就真正建造了直觉主义意义下的展形实数连续统。展形是直觉主义数学中的一个抽象概念，有它广泛的普遍性，Heyting指出：“展形是用一个确定对选择的限制的规则来定义的。”所以如上之展形实数连续统只是借助于展形概念所建造的一个具体展形的实例。而直觉主义连续统一旦建成，就完全改变了古典分析实数论的面貌，直觉主义分析学就能在它的基础上建造起来了。

最后，让我们略提人们对于直觉主义数学的一般评论。普遍认为直觉主义数学所坚持的方法，本来也导源于希望借此排除悖

论的考虑，但是限制过大，只承认一部分最保险的数学，被抛弃的合理因素太多。但联系到计算机数学的发展，构造论方法却有重要意义。详言之，有如下几点评述性意见：

（一）能行性问题具有十分重大的现实意义。正如大家所公认的在使用电子计算机时，尤其不能不注意能行性。

（二）直觉主义数学对于非构造性数学和传统逻辑的绝对排斥是错误的，这种绝对排斥无法解释后者在一定范围内的应用上的有效性。在这一点上，直觉主义派理所当然地遭到了绝大多数数学家的反对。

（三）直觉主义派对实无限性概念的绝对排斥也是不符合科学认识论原则的。

（四）“直觉主义因反对古典逻辑，从而需要把整个逻辑及数学全盘改造，连人们日常认为最简单、最明白无讹的部分也需重新审查，这显然是一件非常艰巨的工程，再由于直觉主义逻辑强调能行性，反变得噜嗦、不方便起来，从而这个数学改造运动极慢，几乎可以肯定难以成功。”（参阅莫绍揆著《数理逻辑初步》1980年版）

§4　略论形式公理学派的观点和主张

长期以来，数学界习惯于把希尔伯特奉为形式主义流派的祖师，往年我们也是跟着那样说的。实际上，那是一种历史的误解。形式主义派的主张和希尔伯特的数学观并不完全相同。现代形式主义者Curry曾明确指出：“有许多人把形式主义与应称之为希尔伯特主义互相等同起来，这是不对的。”（有关详细论述请参考胡世华的文章《数理哲学中的形式主义和柏拉图主义》）

固然希尔伯特确曾有过片面强调形式的倾向，例如他曾说过：“数学思考的对象就是符号本身，符号是这个思考的本质，它们不再代替理想化的物理对象。”但正如Kreisel在分析“希尔

伯特规划”的一文中所指出，希尔伯特决不是一个狂热的、彻底的形式主义者。

我们应该把希尔伯特称之为“形式公理学派”的创始人，以与“现代形式主义者”这一概念相区别。下面我们专来讨论形式公理学派的观点与主张，这些观点与主张称为形式公理主义或希尔伯特主义。兹分条列述如下：

（一）就“无穷观”问题而言，形式公理主义派的观点是认为古典数学中那些包含着“绝对无穷”（实无限）概念的命题确实是“超越人们直观性证据之外”的东西。在希尔伯特与Bernays的1934年的著作中曾表示过这样的意见：“真实无穷乃是通过人们心智过程被插入或外推出来的概念……”但是他们并不同意直觉主义者由于这样的理由去放弃古典数学，包括康托集合论。

（二）既然肯定了实无限概念，也就承认了超穷集合的概念。例如，他们承认全体自然数作成一个完成了的无穷集合。因此无论就有限论域或无限论域而言，他们都主张经典逻辑里的“排中律”是普遍有效的。希尔伯特甚至说过：“数学家使用的排中律就象天文学家手中的望远镜那样重要，是万万不能丢弃的。”

（三）他们主张对古典数学应作形式化的奠定，使之成为形式公理化理论，而这个理论本身必须证明是协调的（无矛盾的）。详细点说，所谓一个数学理论的形式公理化，就是要纯化掉数学对象的一切与形式无关的内容和解释，使数学能从一组公理出发，构成一个纯形式的演绎系统。在这个系统中那些作为出发点而不加以证明的命题就称为公理或基本假设，而其余一切命题或定理都能遵循某些假定的形式规则与符号逻辑法则逐个地推演出来。所谓公理系统的协调性（相容性），指的是这个演绎系统中不能同时包含一个命题和它的否命题。

通常把实现希尔伯特主义的方案叫作希尔伯特规划，其基本内容有以下几点：

（1）证明古典数学的每个分支都可公理化。

（2）证明每一个这样的系统都是完全的，即任意一个系统内的可表命题均可在系统内得到判定（即判定其为真或假）。

（3）证明每个这样的系统都是协调的。

（4）证明每一个这样的系统所相应的模型都是同构的。

（5）寻找这样的一种方法，借助于它，可以在有限步骤内判定任一命题的可证明性。

（四）他们认为验证形式公理化理论协调性所需要的“模型”不能取自感性世界或物理世界。他们提出了以“命题证明法”作为研究对象的一门数学来直接处理公理化的协调性问题。这一门数学就叫做“元数学”或“证明论”。关于元数学的要求是：对形式系统的讨论规定采用构造论的方法，即对推理规则限定用“有穷主义”的方式，不得牵涉无穷的集合概念。（就这一点来说，其方法精神是同直觉主义者的主张相通的）

综上所述，可知希尔伯特主义的基本主张是：一方面希望保存古典数学的基本概念和经典逻辑的推理原则，特别是那些与实无限性有关的概念和方法，诸如无穷集合概念和排中律在无限论域上的使用等等。但在另一方面，又出于可信性考虑，他们几乎和直觉主义者同样地认为可信性只存在于有限之中，而无限性概念不过是理性规定而已。因此他们又把有穷主义的观点贯彻在他们的元数学的推理规则之中。

事实上希尔伯特为了要在有穷主义可信性的准则下保存实无限观点下的古典数学与经典逻辑推理规则，不得不把全部数学划分为“真实数学”和“理想数学”两大类，凡涉及实无限概念和超穷推理方法的数学都称为理想数学。希尔伯特规划就是希望通过有穷主义的构造性方法（包括递归式方法）在元理论研究中证明理想数学的协调性和完备性，以期表明实无限性的理想化成分在应用上的有效性与上述有限性观点获得统一。这就是Kriesel所指出的：“希尔伯特是从有限性的观点出发来理解超穷方法之应

用的。”

但是，希尔伯特主义想把全部数学都纳入到形式公理化宏伟计划中去的企图，已经由哥德尔1931年公布的“不完备性定理”所指明，那是永远不能彻底实现的。

哥德尔的定理是说：“每一个充足的ω无矛盾的算术逻辑L都是不完全的。”我们建议读者最好去阅读M. A. Arbib在1964年出版的《人脑、机器与数学》一书的第五章，从那里可以看到一个最能表明哥德尔定理本质的证明。对该证明经过仔细分析，可以发现这样的事实：不完全性的结论是由于存在一对不能判定的命题（即用“闭型公式”表明的一个命题及其否命题），而这对闭型公式的存在乃是由于存在“非递归可数集”$\overline{Q}$的原故。

因为算术逻辑L是某种“递归逻辑”，而“递归性”是一种潜无限概念，可是非递归可数集$\overline{Q}$的存在性决定于某个完成了的实无限过程，因此，以潜无限进程概念为基础作出来的算术逻辑L之不能完全判定$\overline{Q}$中之全部元素，显然是理所当然之事。其实，哥德尔的不完全性定理无非是下述普遍命题的一种精细的特殊化而已：“总存在某种真无限过程所界定的无穷集合（如$\overline{Q}$等等），其全部内容恒不能由其所相应的任何潜无限进程（如递归性手续等等）所列举或判定。”这个命题可以叫作“无穷过程的层次不可越原理”。

如上所述，哥德尔定理的出现实际上对近代形式公理学派的头上泼了一瓢凉水。公理学派正是因为忽视了实无限与潜无限两者间的本质区别，才错误地认为能用有限主义的方案（如递归方法等）去构造出种种形式系统以表述全部数学真理。

综上所论，可知近代的直觉主义者与公理学派，都因为在无限观上出现了片面性，所以他们设想重建全部数学基础的方案，自然是不可能彻底实现的。

按照反映论的观点，数学理应区分为两大部类，一类是“构造性数学”（即希尔伯特所说的真实数学），另一类是“描述

性数学”（即希尔伯特所谓的“理想数学”）．前者不与实无限过程打交道，故可按直觉主义者的要求去发展；但后者论述真无限过程，就不能采取有穷主义的观点去行事。这里还必须指出的是，数学的抽象形式表现力是有限度的。抽象概念本身总是带有某种片面性和僵化性，所以数学理论的发展绝不能自封于抽象的形式框架里。事实上，数学的真理性并不存在于形式演绎系统的严格证明里，而归根结蒂要通过与物质世界相连系的实践过程去验证。

§5　关于三大流派的简短评论

现代的数学发展史已经表明，三大流派对二十世纪的数学进展有着不同程度的推进作用。事实上若干理论数学分支，所以能取得现今的面貌，都是与上述诸流派的积极影响作用分不开的。

但是，应该指出，三个流派有个共同弱点：他们全都忽视数学科学对象内容的客观实在性，不承认数学对象题材的真正源泉是物质世界。例如，他们把反映客观真理的数学曾分别强调为“纯粹心智的构造”、“纯理性思维的产物”、“自由选定的符号语言”等等。总之，在他们看来，数学仿佛是独立于客观物质世界的某个思维王国中的自由乐园。所以尽管他们之间的意见观点有很大分歧，甚至相互抨击，但不过是对修建这块自由乐园各有不同的主张与行动规划而已。

不难列举出三大数学流派在唯心论哲学阵容里的相应支柱和宗师。似乎特别明显的是，公理主义派无疑是康德主义(Kantism)与“唯理论”的信奉者。此派在数学主张上的实际表现是：“理性给自然制定规律，而不是自然界给我们制定规律”．

我们知道，在《唯物主义和经验批判主义》一书的第五章中，列宁曾深刻地揭示了物理学唯心论的根源。看来真是非常凑巧，我们只需借用列宁的语句，即可相应地用以说明数学唯心论的

主要根源：（1）数学上的唯心论是数学的进步本身所产生的。数学的巨大成就，以及对于这种同质的与纯粹的数学对象（这种对象的形式结构规律可以用逻辑来表现）的模写，产生了数学家对于物质世界的遗忘。“数学的实在内容消失了”，只剩下一些形式符号所编列成的逻辑框架了。在新的发展阶段上，采用新的符号逻辑工具（如元数学、递归函数论等），仿佛得到了旧的康德主义的信念：“理性以规律授于自然”。（2）产生数学唯心论的另一原因，是人们知识的相对性原理。这个原理在旧理论遭到危机时期（例如集合论之“悖论”造成的危机以及欧氏几何之“唯一先验形式”神话的破产）以特殊力量强加于数学家们，而且这个原理——如果不懂得辩证法——将必然引导到唯心论。

还应指出的是，虽然三大流派的基本观点互相对立，分歧很大，但是他们对于整理和重建数学系统的实际作法，确实又有相反相成之处。即就现代充分发展了的“元数学”来看，这三大派的方法论观点，就已经明显地表现出互相依赖、互相渗透的特点。例如形式系统的无矛盾性问题导致证明论的研究，而证明论又须借助于递归函数论方法（或“有穷方法”）；可是递归算法这种思想正是直觉主义者的观念。另一方面，直觉主义者从事于构造性数学以及逻辑主义者讨论他们的演绎系统，也都采用了公理学派的方法。诸如此例，无需一一列举。

最后，我们对三大流派总的估价是：他们确实各有所偏，各有所见；固然他们对待数学本体论的见解是不足取的，但在方法论上却各有重要贡献；特别是直觉主义者的方法论和形式公理学派的思想方法，还值得进一步分析、探讨、继承并发展。

第10讲　略论数学发明创造的心智过程

§1　何谓数学上的发明或创造？

数学对象和数学真理具有客观性，所以对待数学上的创造性新成果，有时用发现来代替发明一词。例如人们常说："牛顿和莱布尼茨发现了微积分原理（即微积分基本定理）。"这就是因为原理的存在是客观的。但是人们也常说："他们发明了微积分学。"这话也说得通，因为微积分学中的一系列符号表示方法和计算法则是由他们创造的。这些都带有人为制作的性质，所以也不妨叫作发明。一般说来，技术科学领域中的创新叫作发明，自然科学领域中原理的觅得叫作发现。至于在数学领域中，发现和发明两词就经常混用了。如果按照极端主义者Kronecker的说法："数学上只有自然数是上帝创造的，其余一切都是人为的。"那末数学上的一切创新或发现都可叫作发明了，但是这位直觉主义先驱者的观点是明显地违反反映论原理的，故不足为训。

数学上界定的大量重要概念和引入的许多巧妙构思确实是人脑的能动性产物，故把它们称作发明或创造是很合适的。例如，实现多值复变函数单值化的黎曼面表示法就是一种巧妙的发明。又例如高斯—勒让德的最小二乘法，庞卡莱的非欧几何模型，A.Robinson的非标准分析等等都可叫作发明或创造。

一般说来，凡在数学上创立新概念、新理论、新模型，提出新方法，证明新定理等，都可叫作数学领域中的发明或创造。

粗略说来，数学上的发明创造有两类，一类是首创性的或开拓性的发明创造，即基本上不依赖于或较少依赖于既有成果的某

种开拓新领域的工作。例如解析几何的发明、微积分的发明、群论的创始、Kummer理想数论的提出以及前一段中提到过的若干例子都属于这一类。另一类是继承性的创造。这主要是指那些在前人已经建立了的重要成果基础上所作的发展性或改进性的工作。例如，现代数学文献上所见到的绝大部分工作成果都属于这一类。当然，发展性的工作中有时也会出现阶段性的飞跃或突破，这又必然带有首创性的因素。所以关于创造性成果的分类，只能有一条模糊的界线。

§2　庞卡莱关于数学创造的论点

庞卡莱（Poincaré 1854—1912）是位举世闻名的博学多产的数学家和物理学家，他的思想活泼而丰富。他曾以自己发现福克斯函数的经历为例表述过数学创造的心智活动规律。其实他所论述的也是一般科学创造的心智过程。他有如下一些论点：

（一）无论是数学还是物理科学，其发明或发现的方法都是相似的。所谓发现或发明无非就是一种"选择"而已。正象在物理科学领域中，选择"可发现定律之事实"，乃是完成各项发现的重要关键那样，数学的发明就是要在数学事物的无穷无尽的组合之中，选择出有用的组合，抛弃无用的组合，从而取得有用的新成果。庞卡莱曾形象地把存在于人脑中的种种数学思想或概念叫作"观念原子"。它们都是一群原来挂在墙上的带钩子的原子。在开动脑子机器后，成群的观念原子便在空中翩翩起舞。原子间的相互组合将能产生新的观念原子。但是组合方式是无穷无尽的，只有通过某种美妙的选择形成的结合（组合）才能产生出极为有用的新观念原子，也即形成数学上有用的新思想或新概念。

（二）选择能力决定于数学直觉。试问：人脑为什么能够将表面上看来并无联系的一对观念原子结合起来产生一个新而有用的概念呢？原因就在于人们头脑里存在着一种关于数学秩序的直

觉，也即关于数学事物关系和谐性的直觉。这种所谓的“数学直觉”正是赖以对无穷无尽的组合（或观念原子的结合方式）中作出有用选择的一种鉴别能力。

庞卡莱认为，有些人具有极好的记忆力，但是上述的数学直觉力不强，所以虽然能够学会和掌握数学，但却无力创造。相反，另一些人虽然记忆力并不极佳，但却具有很强的数学直觉力，因此，这种人能有所发现和创造。可以说，一个人的直觉力的多寡将决定他创造成绩的大小。

阿达玛(Hadamard 1865—1963) 曾发展了庞卡莱的学说，在他所著《数学领域的发明心理学》一书中（有法文本和英文本），曾详尽地论述了选择能力的基础——“数学直觉”的心理学要素。阿达玛认为数学直觉的本质就是某种“美的意识”或“美感”。其实这就是对于数学事物间存在着的某种隐微的和谐性关系与秩序的直觉意识。这种美的意识力越强，发现和辨认隐微的和谐关系的直觉力也就越强，从而选择能力也就越强。

（三）数学直觉导致“最佳选择”的心智活动形式为顿悟，而顿悟产生前存在着一个未被清楚地意识到的“无意识过程”。无意识的活动过程是受着美的意识支配的，它是酝酿着最佳选择或美妙选择的契机，所以无意识活动或大脑的不自觉工作终将导致突然的彻悟即顿悟的出现。

但是无意识过程往往产生于长久的自觉工作之后。假如没有自觉的工作阶段，即缺少有意识地自觉地开动大脑机器的阶段，则悬挂在墙壁上的那些有钩子的“观念原子”，就不能被动员起来使之翩翩起舞；从而也就谈不到有任何相异的“观念原子”互相结合的机会，当然也就更不会有什么组合的选择可言。

总之，若是事先没有自觉的工作和有意识的努力，则大脑中也不会有无意识的活动和不自觉的工作，因而也不会有顿悟出现。

那末为什么无意识活动过程能产生有用的顿悟（即观念组合的最优选择）呢？这是因为对于事物关系和谐性的美感（或美的

意识）能力在无意识的心智状态下可以免除任何条条框框的限制，能最自由地、从容不迫地去作出最优选择的原故。

不只是数学创造，即使象文学艺术领域的创造性灵感或顿悟，也往往产生于无意识活动阶段。例如中国宋代大文豪欧阳修谈到过所谓“三上文章”的经验，他的一些文学作品的构思往往产生于厕上、马上和枕上。这就是说，正是在厕上、马上或枕上的时候，特别有利于欧阳修的文学思维进入无意识阶段，从而能产生出美妙的文思或灵感，类似于科学研究中的顿悟。事实上，枕上能形成科学发明的顿悟现象，科学史上也不乏其例，例如笛卡尔解析几何的萌芽思想即产生于凌晨枕上初醒的时候。

很显然，庞卡莱与阿达玛关于数学创造的心智活动过程中的“无意识活动”的理论，乃是对心理学上顿悟形式的深入阐发。他们的有关论述，对现代心理学的研究工作者有相当影响。他们的一系列论点多少阐明了自觉工作与不自觉的工作、有意识的努力与无意识活动之间的辩证关系，应该说，在科学哲学上是有其合理性的一面。

往年有些人往往把庞卡莱—阿达玛关于数学创造的心智活动学说，当作是纯粹唯心论的东西看待，因而不敢多作介绍。看来这是一种不必要的误会。

§3　略谈数学创造的一般心智过程

上一节中谈到了有意识工作的重要性。用庞卡莱的话来说，只有自觉的有意识的工作才能驱使“观念原子”飞舞起来，并通过观念原子间的结合产生出新的思想或有用概念。若按现代创造发明学的观点来说，就是要积极开动大脑机器，掀起所谓“脑风暴”。

脑风暴常常出现在心理学上所谓的“烘热期”。那是一种在脑海中迅猛地涌现出种种现象、联想、猜想、假设和非逻辑思维的心智活动形态。

人脑思维运动的烘热期或脑风暴不能立即产生，一般至少需要精神贯注地连续工作一两小时或两三小时才能逐步形成。

有效的脑风暴必须连系着一个明确的目标——如需要解决某个数学问题，或希望取得某种数学发现等目标。一般说来，开展脑风暴的目的性越强，其效果便越好。这是十分显然的，脑风暴所连系的目标或目的，将成为思维运动的联系中心，于是脑海中涌现出来的联想、猜想、假设和一切非逻辑思维将会围绕该中心展开，从而有利于解决问题达到目标。

根据“发明心理学”的研究，在脑风暴的初始阶段通常只能涌现出一系列十分平凡的思想、观念及其组合，而这些显而易见的观念及其组合往往并不能帮助解决问题。因此要想获得创造或发现，还必须使脑风暴坚持进行下去。比方说，使脑风暴持续两个小时或一个半小时，则最后半小时就可能出现寥寥无几的或极为罕见的新思想或新观念，可是它们却是较为深刻的甚至可能是有助于解决问题的思想或观念。

根据庞卡莱和阿达玛的学说，彻底解决问题的新思想或新观念（或所谓顿悟）通常未必出现在脑风暴进行之际，而往往产生在脑风暴平息之后的无意识活动阶段。上一节中已经详细地阐明了无意识过程对选择观念组合的特殊重要作用（无意识过程或无意识境界并未让一切观念原子都停止了活动；而通过审美意识选择出来的个别观念原子间的奇妙组合却会给出有用的顿悟）。

一般说来，顿悟只是帮助数学研究工作者形成合理的设想或猜测，而要实现设想或验证猜想还须继续进行有意识的工作甚至是艰巨的工作。因此要完成数学上的任何创造或发明，通常总须从有意识的工作开始，最后还须用有意识的工作来结尾。

那末，是否任何创造发明过程都必定包含着无意识的活动过程呢？换言之，无意识作用是不是发明创造的必要条件呢？看来未必如此。固然数学上的创造或发明往往联系着新思想、新方法或新概念的提出或引入，而崭新的思想、概念或方法又往往来源

于顿悟；但是顿悟既可能产生于无意识活动阶段，也可能形成于脑风暴进行的末尾。

从事创造性活动的数学工作者，既要善于进行发散思维，又要善于进行收敛思维，发散思维是一种取得合理设想或猜想的思维形式，它包括着联想、想象、模拟、类推和直观推理（又称似真推理）。显然在脑风暴进行过程中，发散思维扮演着主要角色。

因为数学创造可分为取得合理设想和验证合理设想两个基本阶段，而验证主要是靠合理的逻辑思维，也就是收敛思维，所以富于创造力的数学家还必须是擅长于收敛思维的高手。

但是，一般说来，数学上的新思想、新概念和新方法往往来源于发散思维，所以按照现代心理学家的见解，数学家创造能力的大小应和他的发散思维能力成正比。详细说来，任何一位科学家的创造能力可用如下公式来估计：

创造能力＝知识量×发散思维能力

这个公式有一定的道理，因为知识量越大，则联想、类比、想象的领域就越广，从而产生出新思想、新概念和新方法的机会也就越多，所以科学家（包括数学家）的创造能力又同他的知识量成正比。事实上，很难设想知识面极狭窄的学者能有多大发明创造能力。

发散思维是以感觉或直觉为基础的非逻辑思维形态，所以在数学创造性思维过程中，它是一种不严格的思考方式或方法。数学必须要求精确而严格，所以接受过数学教育熏陶的人又往往会排斥或鄙视不严格的发散思维形式。但是，真正具有数学创造力的科研工作者，按照前面介绍的公式来看，就必定是既善于严格思维又善于不严格思维的人。事实上，由前面的详尽分析早已能够得出结论："数学创造往往开始于不严格的发散思维，而继之以严格的逻辑分析思维，即收敛思维"。

仔细说来，数学创造的心智活动过程是异常复杂的。要阐明一些过程细节，非剖析实例不可。但为了节省篇幅，此处仅就一般心智过程略作介绍而已。